A CRITICAL PROJECT

By

T. S. Rowe

A thesis submitted in partial fulfilment of the
requirements for the Degree

of Master of Arts in Philosophy

Department of Philosophy and Religious Studies

University of Canterbury
March, 2008

Acknowledgement

I wish to extend my great appreciation to Derek Browne for his valuable guidance, critical commentary, and help in the preparation of this thesis for examination. In addition I would like to express my gratitude to Allen Buchanan for his encouragement of this project, along with my appreciation of his valuable work in the area of social epistemology, which has in many ways framed and given direction to this thesis.

Contents

Preface

Knowledge according to Francis Bacon is power, and something according to Aristotle that each of us desires. Few of us seriously doubt that having true beliefs is of great importance to our lives and well being. Through the veracity of our beliefs we are able to obtain a level of control over ourselves and the world that is otherwise unavailable to us. Our very lives can often depend upon our beliefs being true, such as when we inquire from another about the way to the hospital in an emergency, or when we simply cross the road. While some falsehoods are harmless and cause no recognisable harm in their being believed, others can directly and indirectly lead to catastrophic results both for ourselves and for others. The belief that witches were in league with the devil and contrived to use dark magic for evil purposes led to the brutal execution of thousands in Europe, while an Aztec belief that the thirst of a ruling god could only be slaked by the blood of youths and maidens led to the sacrifice of a staggering 50,000 people in one year alone, with tens of thousands of others sacrificed annually.[1] In whichever direction we turn it seems that we find ourselves having a great stake in the truth of our beliefs, whether we acknowledge this or not. Critical thinking is widely upheld as an essential tool in our truth-seeking. It involves most centrally the holding of thought and judgment to various regulatory and normative intellectual principles which are taken as conducive towards a) identifying truth and b) limiting or reducing the possession and acceptance of falsehood. Where truth is a doxastic goal, understanding in what ways and to what extent our veridical intentions are compromised is of significant value.

It is with a view to this that this thesis examines what are for us two great sources and causes of error. The first of these examined arises from the influence of cognitive biases upon our thinking. Treated in the literature as divisible into 'hot' and 'cold' categories,[2] the foremost amongst these describes influences brought about by the effects of subjective preferences, motivations, emotions, and interests, which may motivate us to satisfy those elements at the cost of truth. 'Cold' biases, in contrast, are unimpassioned influences on reasoning that distort it in distinct, preference-extrinsic patterns. The second great area of concern for truth emerges as a result of our wide-ranging dependence upon others for a vast amount of our beliefs about the world, and how through this dependence we are opened up to the possibility of believing many false things on the basis of our trust in others. Two critical questions are asked: a) in what ways can experts arrive at false beliefs; and on the basis of the epistemic division of labour, b) in what ways can these false beliefs be transmitted and believed more widely.

---

[1] H. L. Mencken (2000), pp. 43-45.

[2] Alvin Goldman (1999), p. 230. It should be noted that there are more biases that influence our thinking than can be catalogued here, given constraints, and for a more extended and detailed account the reader should refer to the secondary literature on the subject.

The fact that we can fall into error as a product of our social relations is ever more important in light of the many ways in which we can believe erroneously ourselves, for it might be argued that in view of our cognitive imperfections the best way forward is for us to place an ever-greater weight of responsibility in trusted domains of expertise to provide us with true beliefs. By examining the historical, psychological, and linguistic characters of epistemic expertise and relationships, I will attempt to persuade us away from viewing this option as attractive by drawing attention to the ways in which domains of inquiry may not simply fail to rise above many of the biases found widely, but may also be burdened with truth-undermining features particular to the very status of expertise itself. Radical scepticism in relation to epistemic authorities will be avoided; instead it will be argued that a healthy and fruitful epistemic situation is both possible and preferable in the form of what Allen Buchanan has called 'epistemic egalitarianism'[3] – a situation wherein epistemic authority is approached in a moderated, critical fashion.

Later on discussion will turn towards various conceptions of the justification of testimony with an aim of further elucidating exactly what sorts of things, if any, can be gleaned about epistemic relationships by looking at the nature of testimony itself. It will be argued that certain conceptions of the justification of testimony lend themselves naturally towards a critical attitude in relation to the testimony of kinds of expertise, and that some propositions are more favourable than others as candidate items of deferral. The final chapter examines some approaches to irrationality that have been advocated. It is argued that there is significant value in steps being taken in order to educationally equip people with the intellectual skills and background knowledge necessary to raise their capacity for and willingness to engage in critical thinking. For ongoing discussion it will be assumed that a general principle of uniformly seeking truth over falsehood provides for the greatest possibility for the most utility overall.[4]

[3] Allen Buchanan (2004)

[4] Philip Kitcher's (2001) has interesting discussion of the way in which we need not see truth as in all cases utility increasing. And indeed it could be argued that some cognitive biases have a greater instrumental than veridical value (such as allowing the subject to think more positively of him or herself beyond what the evidence can warrant), and as such that some falsehoods need not or even *ought* not be rectified.

A Tabling of Bias

2.1 The Pressure Towards Conformity of Belief With Others

Experiments published in 1951 by Solomon Asch threw light upon the pressures we face in regards to believing as others do. Asch's simple experiment saw individual subjects placed in a room with seven others and asked to specify which of three lines drawn was equal in length to a fourth. Unknown to the study participants was that the seven others in the room were collaborators in the experiment, and all would provide incorrect answers.[5] Due to ordering the subject would answer last, meaning that if they were to give the right answer he or she would have to do so against the stated judgments of every other person in the room. As it would turn out, this proved to be a monumental task for many people.

Despite involving *very* simple cognitive judgments about the relative lengths of lines, a staggering three quarters of people were found to give the wrong answer at least once during the experiment, with many doing so multiple times. Only one quarter of people provided correct answers every time. When the lines were very close in length to one another it was found that nearly everyone succumbed to the popular choice. Significantly, of those who did give correct answers, most did so showing obvious signs of nervousness and hesitancy.[6] This was taken to indicate that those who exhibited that response were not simply a small group of people uncommonly impervious to the pressure of conformity. Many subjects reported being unaware that they were subject to the influence of those around them in relation to their stated perceptual judgments, which they believed to be independently arrived at.

Subsequent experimentation found that if just one of the collaborators present in the room gave the correct answer before the subject gave theirs, the rates of conformity decreased noticeably. Also found was that if subjects wrote their answers instead of registering them verbally then the incidence of conformity reduced.[7] This suggests that the pressure we face to accord with the judgment of others relates both to how salient our dissent will be in the eyes of others, and what degree of consensus exists on the matter under question. The experiment was repeated in different formats and using subjects from different cultural backgrounds with similar results. The underlying bias evidenced here has deep significance for political and social arrangements, ethics, and philosophical conceptions of individuality and autonomy, and provokes questions about even the most basic realities of our social existence. It illustrates how the putative judgments of others can be, under the appropriate conditions, cognitively powerful enough to divest us of our most basic sorts of judgments - or at a minimum, prevent us from correctly reporting them. Other experiments have confirmed this conformity tendency, but have also

[5] Stuart Sutherland (1992), pp. 45-46.
[6] Ibid.
[7] Ibid, p. 47.

suggested that under the right conditions initially strongly biased reporting, once freed from the influence of sources of deviation, can right itself over a period of time.[8]

2.2 Obedience to Authority

> With numbing regularity good people were seen to knuckle under the demands of authority and perform actions that were callous and severe. Men who are in everyday life responsible and decent were seduced by the trappings of authority, by the control of their perceptions, and by the uncritical acceptance of the experimenter's definition of the situation, into performing harsh acts. A substantial proportion of people do what they are told to do, irrespective of the content of the act and without limitations of conscience, so long as they perceive that the command comes from a legitimate authority.
> --Stanley Milgram.[9]

Probably the quintessential example of the way in which we can be influenced by the demands of authority is Stanley Milgram's 1960's experiments on obedience. In Milgram's study test subjects believed they were partaking in an experiment to see how punishment affected learning.[10] Participants were set up so as to believe that they would be electrically shocking another person (who was out of sight, but with whom there was audio contact) for incorrectly answering word association questions.[11] The person supposedly being tested may have for example been asked to pair the word 'box' with the word 'blue', and then during examination would have been given four possible words from which to correctly identify the earlier correlated pair. The punishment for incorrectly answering was the fictitious application of an incrementally increasing electric shock. In fact, no real shocks were administered, but screams were recorded on a tape and played to the test subject in order to create the perception that the experiment and procedure was authentic. The shock dial read from 15 to 450 volts, with descriptions of the level of pain associated with the voltage, such as 'slight shock', 'strong shock', and 'danger: severe shock'. For each wrong answer the dial would be levelled up 15 volts. The subject was told that although the shocks could be extremely painful, no permanent harm would result.

The study supervisor, described as a 31 year old stern man who wore a grey technician's coat so as to present an authoritative figure, was placed in the same room as the test subjects. This proximity to the test subject would allow for the greatest inclination towards being obedient. Having complained earlier of heart problems, the collaborator would appear to remain silent after a shock of 330 volts was delivered, in order to give the impression of having been killed by the stress of the testing procedure (or at least as having fallen unconscious). With silence as the only response to the questions at that point, the subject would be told to *continue to apply* the electric shocks, as failing to give an answer was to be considered as equivalent to failing to give an answer correctly.

---

[8] Peter J. Richardson and Robert Boyd (2005), pp. 122-123.
[9] Thomas Blass, 'StanleyMilgram.com'.
[10] See Stanley Milgram (1963), pp. 371-378 for the cover story given and a full description of the methodology.
[11] Stuart Sutherland (1992), p. 36.

The results of the study profoundly unsettled those involved, and were as follows: in the first experiment, 25 out of 40 people continued delivering shocks to the maximum level – 450 volts. Nobody stopped before the 'moderate' shock level was reached, and several went to 'extreme intensity shock'.[12] The experiment was run several times with results that proved to be equally as dismaying. As many women were found to give the maximum shock level as men.[13] It was found that if the subject was in the same room as the person feigning electrocution, the number of those who went to the maximum level dropped to 1 in 4. Even without the technician in the room to provide an authoritative physical presence, 9 people in 40 proceeded to the highest possible level. Many of the subjects gave indications of substantial psychological stress in completing their task, including sweating, trembling, stuttering, biting their lip, clawing their finger nails into their skin, and pleading to be 'allowed' to stop, all of which was taken as evidence for the fact that they did not continue the testing procedure on the basis of believing that the experiment was not genuine. The supervisor responded to requests for the cessation of testing only by urging subjects in firm language to continue ("please continue", "you must continue", etc). Nervous laughter was also common, which was described later not as having a sadistic origin, but rather as part of a coping mechanism induced by the stress of the experimental procedure.[14] It was noted that some subjects displayed no concern regarding the suffering of the patients, and maintained a quiet satisfaction at having completed their task to the requirements of the test technician.[15]

At the completion of the testing phase subjects were debriefed on the true nature of the study. Although at this time they were then availed of an opportunity to save face by claiming that they had known all along that the shocks administered were fake, almost none chose to do so. Milgram reported instead that many expressed being profoundly affected by the experiments in positive ways. One remarked that "this experiment has strengthened my belief that man should avoid harm to his fellow men even at the risk of violating authority", a reputedly typical point of view held by test subjects.[16] The experiments were repeated using people from different cultural backgrounds, with similar results.

Upon publication Milgram drew intense criticism concerning whether his texts were ethical. In his own defence Milgram noted that many of the subjects, seeing themselves the merits of the experimentation, had proactively contacted Milgram to inquire into the possibility of volunteering for involvement in future experimentation. The experiments stand as one of the most important and unsettling insights into human psychology we have, vividly revealing how in the right circumstances our desire to be compliant with authority can overwhelm an individual's independent moral sensibility. The phenomenon uncovered holds the potential to play an explanatory role in a number of the ethically tragic historic events.

---

[12] Ibid, p. 36.
[13] Sutherland notes that while women may on one hand be more compassionate, they may also be more prone towards being obedient to authority, and hence may have been more influenced by the authority figure.
[14] Stanley Milgram (1963), pp. 371-378.
[15] Stuart Sutherland (1992), pp. 37-38.
[16] Ibid, p. 38.

2.3 Public Commitment and the 'Boomerang' Effect

> … human frailty cannot endure even the most just refutation of an error, unless it is tempered by soothing and flattery, and hardly even then…
> --Arthur Schopenhauer, *The World as Will and Representation*.

One study conducted purported to show an inherent irrationality involved with the public admission of our commitment to matters of some value.[17] Of a group of women living at Yale who strongly favoured the disseminating of information regarding birth control, half were asked to sign a public petition endorsing birth control education at a local public high school. The next day a persuasive leaflet against such an action was given to half of those who had signed the petition the previous day, and half of those who had not. The leaflet for example argued that it would promote promiscuity, and that this sort of information should be given to children by parents. While all of the women were initially strongly in favour of public information regarding birth control, there were now four distinct groups of women categorisable according to who among them had signed the petition and who had seen the pamphlet:

1. Women who had signed the public petition to inform school children about birth control, and who had later seen solid argumentation made against it.
2. Women who had signed a public petition to publicly inform school children about birth control, but who had not seen this solid argumentation made against it.
3. Women who had *not* signed a public petition, but who *had* later seen solid argumentation made against it.
4. Women who had *not* signed a public petition, and who had *not* seen solid argumentation made against it.

The women were then all asked a series of questions in order to find out how many would volunteer to participate in work aimed towards positively disseminating that information. It was found that the women who had *not* signed the petition, and thus had *not* made a public committal in favour of spreading information about birth control, and who *had* read the pamphlet arguing against what they had initially supported, were less likely to agree to volunteer to disseminate that information on birth control than those who had not received the pamphlet but who had not signed any petition. This showed that they were responsive to the arguments made against their position on the matter in rationally consistent ways.

However, the opposite effect was found to obtain for those who *had* signed the public petition and *had* seen the pamphlet: more of them agreed to volunteer than those who had not seen the pamphlet but who had also signed the public petition. So in other words, for those who had publicly taken an affirmative position on birth control education, hearing solid argumentation against this only served to entrench and invigorate their views, where for others who had not publicly made commitments, it weakened it. Good argument

[17] Stuart Sutherland (1992), pp. 49-50.

contrary to their views in effect 'boomeranged'. This of course is objectively irrational – our judgments of the strengths of cases should not depend upon whether we have voiced our opinion publicly in relation to it, or kept silent. The psychological explanation offered for this phenomenon is that those who had made no public statement were 'freer' to change their minds without any negative social impact, whereas those who had taken a position publicly were more committed to their views through having endorsed them in a socially visible way. As a result of that endorsement, they felt themselves more compelled to justify and defend that position afterwards.[18] And if there are compelling arguments against ones own position on a matter, to the extent that the argument is compelling or challenging we may seek to be equally adamant of the strengths of our own position.

This bias has substantial implications for the way in which we think about the nature of debate and reason in free societies. We are likely to find it harder to change our minds about an issue if we have publicly endorsed or condemned it, and contrary argument and evidence may (through this) only serve to irrationally strengthen our commitment to a position. To give but one possible outcome, politicians (whose statements may be the most public of all) may find themselves in a very hard psychological conundrum when challenged to change their position or reverse policies which they have previously supported. It would accordingly serve us all well to internally check whether our motivations for supporting some position genuinely derive from the strength of that position as seen, and not partly or mostly from our public history with that position.

2.4 The Power of Suggestion and Phraseology

The structure of the way in which questions are placed to us can affect judgments we make. We have a tendency to select options in the middle of ranges. If for example a group of people was asked to choose a random number between 1 and 10, statistically more choices would fall somewhere within the middle of that range than would be evenly distributed across its entirety, or inclined towards to the ends. Consider a hypothetical questionnaire surveying political satisfaction for a political figure laid out in the following manner:

Dissatisfied Satisfied Very Satisfied Extremely Satisfied

The given format biases answers to be given as more satisfactory than not, given that positive answers unevenly occupy the middle territory.[19]

More significantly, two aspects of a study by Loftus and Palmer show how suggestion can influence us with respect to memory.[20] In the first part of the study students were asked to watch videos of two cars crashing, and answer questions about what they had just seen. One of the questions asked how fast the cars were travelling when they […],

[18] Ibid.
[19] Stuart Sutherland (1992), p. 231.
[20] E. F. Loftus & J. C. Palmer (1974), pp. 585-589.

where one of five descriptive words was used: contacted, hit, bumped, collided, and smashed. The students who were asked the more provocatively termed question (i.e. 'smashed') subsequently estimated the cars to have been travelling at faster speeds than those asked with the more moderate terms employed (e.g. 'contacted', or 'hit'). The second part of the experiment occurred one week later. Students were contacted and asked whether they recalled any broken glass in the videos. Despite the fact that there was *no* broken glass visible in the videos, subjects who were asked questions using the stronger language were found to be much more likely to recall having seen broken glass. Such research supports models of memory as being reconstructed rather than 'recorded', and shows how the influence of language can extend far beyond what judgments we come to, but to even what we remember. Our susceptibility to political and social manipulation through language was famously explored by George Orwell in his novel '1984', but outside of narrative the influence of language in public affairs is potentially sweeping. Lawyers, for instance, may be found to use the power of suggestibility to influence jurists, and through nothing more than their choice of words push forward an interpretation of events in a direction that is to their advantage. To give a contemporary example, savvy politicians have found it prudent to label tax cut plans not 'tax cuts' but 'tax relief', as the word 'relief' colours our impression of the idea with connotations of ease and release. This is conceptually preferable, as our brains are conceptually primed with the *avoidance of loss* (as 'tax relief' implies) carrying greater cognitive significance than *gain* (as 'tax cut' implies).

2.5 Loss and Risk Aversion

One study asked people to choose between the following:

> Situation 1
> Option A: Accept $50 with certainty.
> Option B: Accept a gain of $100 with a .5 probability.
>
> Situation 2
> Option A: Accept a certain loss of $50.
> Option B: Accept a loss of $100 with a .5 probability, and a .5 probability of losing nothing.

It found that in situation 1 most people chose option A, whereas in situation 2 most people chose option B.[21] (the aforementioned tax cut/relief conceptual asymmetry likely depends upon this effect). Loss may be disliked as much as twice as much as gains are favoured, with a loss of $100 creating twice as much disutility for us as a gain of $100 creates utility.[22]

Paul Rubin notes a further study in which it was shown that most people refused a 50-50 bet in which they could either lose $10 or gain $11.[23] It is always statistically rational for any person to take such a bet, since the bet is favourable (and would most likely come out as such in a sufficiently long run).

[21] Stuart Sutherland (1992), p. 223.
[22] Paul Rubin (2002), p. 173.
[23] Ibid, citing Matthew Rabin (1998), pp. 11-47.

2.6 Confirmation and Disconfirmation Bias

> These theories [of Marx, Freud, and Adler] appeared to be able to explain practically everything that happened within the fields to which they referred. The study of any of them seemed to have the effect of an intellectual conversion or revelation, opening your eyes to a new truth hidden from those not yet initiated. Once your eyes were thus opened you saw confirming instances everywhere: the world was full of *verifications* of the theory.
> --Karl Popper, 'Science: Conjectures and Refutations'.

Confirmation bias describes a tendency for us to place an over reliance on instances that confirm the existence of a relationship.[24] We tend to orientate towards positive, confirming instances of hypotheses rather than towards negative, disconfirming instances or evidence. Gilovich offers the example of a certain belief some people have that couples who are having difficulties in conceiving a child are more likely to do so after having adopted. The underlying explanation for this seems to be that stress is thought to contribute to their inability to conceive, and with an adoption of a child their stress levels are thought to be lowered though having the pressure to conceive in turn lowered. So relieved, future conception is then thought to be more likely. In examination of the evidence, holders of this belief will find that examples of couples who conceived after an adoption appear to them to be far more salient evidentially than couples who did not conceive. Such examples then appear to confirm the underlying explanatory theory, and so validate their beliefs. In fact, the belief in question however is not supported empirically, as adoption has not been found to lead to any statistical increase in couple's ability to conceive.

The satisfaction of preferences can partly account for the prevalence of confirmation bias. Given a preference for some *x* being true, we may find ourselves naturally inclined towards giving significant weight to evidence that is confirmatory of *x*, while ignoring or explaining away evidence that is disconfirming of *x*. When we find that initially obtained evidence supports our preferences, we are generally satisfied to end our analysis, having found support for our beliefs. However, when we find initial evidence is hostile to our preferences, we often continue our evidential search in order to seek out further supporting evidence, or evidence that can be used to undermine that which is hostile to the theory we favour. In such a way our chances of finding support for preference-satisfying beliefs are dramatically increased. There are limitations however to our ability to do this, for in trying to arrive at preferable conclusions we also generally try to be rational by having justification that could persuade dispassionate observers. Thus to a certain extent reality constrains our freedom of belief; that is, our freedom to have perfectly preference-satisfying beliefs. Concerns for objectivity can, despite the biases that slant belief in our favour, force us to "acknowledge and accept undesirable conclusions".[25]

[24] Thomas Gilovich (1991), p. 31.
[25] Alvin Goldman (1999), p. 236, citing Ziva Kunda (1990), pp. 482-3.

Our inclination however towards confirmation over disconfirmation cannot be explained purely in terms of the satisfaction of preferences, as there is certain evidence to suggest that we are orientated towards confirmation dispassionately, independently of any obvious desires for particular outcomes, such as in the case of simple introduced numerical rules.[26] An example to illustrate this is given by the case of students who were allocated 20 questions with which to identify a number between 1 and 10,000. When they would ask whether it was between 5,000 and 10,000 and the answer was in the affirmative, they would cheer, but when the answer given was negative they would groan. Their respective joy and disappointment for the answers that they received was misplaced, since in either case the answer would have been equally informative, narrowing the margin by essentially equal degrees.[27]

The effects of confirmation bias have significant consequences for belief preservation. In a study of student psychotherapists it was found that once they had accepted some diagnosis they could look through an entire folder of contrary evidence and interpret it in favour of what they had accepted. This would tend to indicate that once accepted, our beliefs can be somewhat resilient to falsification.[28]

Confirmation bias's influence may be partly due to the fundamental operation of our neurology and its means of handling of information. It seems that positive statements are just simply easier to deal with cognitively than negative statements. A simple example illustrates this point: consider the proposition "All Greeks are mortal" and compare it to the proposition "All non-mortals are non-Greeks".[29] While both mean precisely the same thing, the former is clearly much easier to deal with than the latter is.[30]

The flipside of confirmation bias is a disposition towards seeking the disconfirmation of evidence and theories which contradict, and thus compete with, our own beliefs. We tend to spend more time in critical evaluation of theories that oppose our beliefs than we do with theories that we agree with.[31] This added time is spent seeking grounds on which to undermine conflicting evidence in order to discount it, and so preserve our pre-existing beliefs. We can see much of this in action in a study in which brain scans were conducted on staunchly pro-democrat and pro-republican subjects who were tasked with resolving the recorded contradictory statements of 2004 U.S. presidential nominees George Bush and John Kerry. The study revealed that the part of the brain most closely associated with reasoning (the dorsolateral pre-frontal cortex) remained quiet, while the parts of their brain associated with the processing of emotions, conflict resolution, and judgments concerning moral accountability (the orbital frontal cortex, anterior cingulate, and posterior cingulate respectively) had heightened activity. Once a preferable judgment concerning the contradictions was arrived at, heightened activity in the ventral striatum

[26] Thomas Gilovich (1991), p. 33.
[27] Ibid.
[28] Theodore Schick, Jr. and Lewis Vaughn (2005), citing L. J. Chapman and J. P. Chapman (1967), pp. 193-204.
[29] Schick and Vaughn (2005), pp. 31-32.
[30] This may help explain why Hempel's 'Raven Paradox' causes us intuitive difficulty.
[31] K. Edwards & E. E. Smith (1996), pp. 5-24.

was observed – a part of the brain linked to pleasure and reward.[32] This would tend to indicate that the confirmation of our beliefs and the disconfirmation of challenging beliefs is bound up in pleasure-seeking and pain-avoiding motivations.[33]

Incorporating both confirmation and disconfirmation tendencies, a picture of our unenlightened cognitive position develops which shows us as 1) generally seeking to avoid exposing ourselves to evidence that might contradict and disprove our views, 2) often refusing to believe such evidence where we come into contact with it, 3) allowing existing beliefs to distort new evidence so as to make it consistent with our beliefs, 4) selectively remembering items according to whether they are in line with our beliefs or not, and 5) and seeking to maintain our self esteem by avoiding having to consider ourselves wrong on an issue.[34] All this of course is worrying for conceptions of ourselves as rational beings, even given motivations to establish the justifications for our beliefs on such terms that they would be convincing to impartial observers. Nevertheless, if we are consciously aware that confirmation bias is a factor that plays upon our thinking then self-monitoring of our thought processes can come into play and offer the possibility of reducing their impact.[35]

2.7 The Danger of De-Individuation

Anonymity can alleviate the burdens of moral judgment and responsibility and allow individuals to act morally vicariously without normal ethical commitments. In one experiment subjects were encouraged to give electric shocks to a study collaborator (who feigned being electrocuted). Some of the subjects were cloaked and hooded, such that their identity was hidden from observers. Subjects disguised in this way gave on average more shocks than those who were not.[36] That a person's moral inhibitions and judgments can be diminished under certain conditions displays an important fact about the contingent nature of human morality. Rioting and soccer hooliganism offer two vivid examples of social phenomena that exhibit this sort of effect, with the latter definitively incorporating in-group/out-group tendencies.

2.8 In-Group and Out-Group Bias

'Man is a social animal', disposed towards forming himself and identifying himself within social groups for his advantage. An 'in-group' is any group to which we belong, while an 'out-group' is by contrast any group to which we do not belong. Drawing

[32] Michael Shermer (2006).
[33] I take it that everyone reading this is sufficiently familiar with the disutility of cognitive dissonance to grant that such a state is psychologically uncomfortable.
[34] Stuart Sutherland (1992), pp. 151-152.
[35] It has argued that in cases it may be not so much that we have a tendency to confirm propositions that explains why we appear biased towards the confirmation of hypotheses, but that we may have a questionable ability to generate alternative hypotheses to stand as hypothetical alternatives (Barbara Koslowski, Erskine Seminar, University of Canterbury, August 4th 2006). This seems plausible, and if true would provide an immediate route for the partial amelioration of confirmation bias through the supplying of plausible alternative hypotheses to those under consideration.
[36] Stuart Sutherland (1992), p. 55.

distinctions of 'we' and 'they' is a fundamental human activity,[37] and all humans classify themselves as part of groups beyond the family unit.[38] Centrally, we classify ourselves into groups based upon distinctions among kinship, sex, and age. We are generally disposed to think of non group members of the most important types as being "alike and as being enemies, while members are treated as individuals".[39] The complexities of group identification are significant however, and can include many subsets. Paul Rubin's use of his own case provides a humorous illustration of this flexibility:[40]

> Within economics, I may identify myself as a microeconomist when the department is considering whether to hire a microeconomist or a macroeconomist. But when the college is considering whether to award a position to a sociologist or an economist, all economists join in decrying sociology. Economists and sociologists may agree to hire a social scientist instead of a humanities professor. Humanities professors and natural and social scientists will all agree that the budget for the College of Arts and Sciences should be larger relative to the professional schools, but we agree with our colleagues in the Medical, Law, and Business Schools that the university needs more money.

Tendencies for grouping are advantageous for individuals insofar as they promote cooperation in the pursuit of shared interests where competition for limited resources exists. Indeed, the division of groups can imply a competitive relationship between groups, a point reflected in the fact that we tend to give preferential treatment to members that we recognise as being part of the group with which we associate.[41]

That we do engage in making such distinctions is not at all trivial, as the consequences of the way in which we divide up the natural and social world are far reaching.[42] For instance, classificatory schemes that divide the human world along racial lines may become so embedded within our way of thinking as to make it difficult for future generations to see outside of them. Such distinctions may become interwoven in an extensive network of legal, political, and social institutions, and aid to foster self-conceptions conducive or inhibitive of diverse actions. Special attention then deserves to be paid to the practice of determining 'we' from 'they', and the grounds that such distinctions are held upon examined for their merit.

2.9 Primacy Effect

[37] H. Tajfel (1970), pp. 96-102
[38] Paul Rubin (2002), p. 34.
[39] Ibid, p. 32.
[40] Ibid, p. 35.
[41] Tajfel, H. (1970). A teaching assistant related an insightful example to me of how she taught a class where, due to the arrangements of seating in the room, the class was divided effectively into two. The students self-organised themselves so that the older students sat on one side of the room and the younger students on the other. As discussions unfolded the class atmosphere quickly deteriorated into an antagonistic one, with both sides becoming adversarial towards the other. Of particular note was the passive-aggressive behaviour displayed by the older students who would roll their eyes at anything the younger students offered in discussion. After a couple of weeks of this behaviour she reproached them and made them push the seating together. With the eradication of the means by which they could delineate themselves into groups, the out-group tension evaporated.
[42] Philip Kitcher (2001), chpt. four.

Primacy effect describes a tendency to allocate greater weight to initially supplied data. One group of people was asked to estimate quickly (i.e. without mentally calculating) the product of:

8 x 7 x 6 x 5 x 4 x 3 x 2 x 1

While another was asked to estimate the product of:

1 x 2 x 3 x 4 x 5 x 6 x 7 x 8

The average answer for the first was 2250, while the second was 512 (the correct answer, for the record, is 40,320). Each group is taken to have been influenced in their estimations by the numbers appearing at the beginning of each sequence.[43] The effect can also been seen in the following: consider a person described as "intelligent, industrious, impulsive, critical, stubborn and envious". Now consider another described as "envious, stubborn, critical, impulsive, industrious and intelligent". In a study conducted by Solomon Asch it was found that people who read the first list were much more likely to give positive estimations of, for instance, the happiness of the person, than those who were given the second.[44] The explanation offered is that words are used to build a mental picture of the person in question, which in turn by their order influence the picture that is built differentially.[45]

2.10 Self-Serving Bias

There is a saying in the art industry that a person's willingness to believe their painting is a forgery is directly related to how much they paid for it. As was seen in some of the earlier examples of confirmation bias, our preferences can act greatly upon our belief forming practices by motivating us to come to certain judgments and beliefs. One way in which this is done is through a selection of cognitive strategies in order to reach desired doxastic goals. More broadly this tendency is known as self-serving bias – our propensity to form beliefs that are in some way beneficial to us.

A number of examples bring to light this feature of human psychology. In one study participants were told to rotate pegs for 20 minutes (a pointless and monotonous task). After completion, they were asked to tell the next study participants that the task was very interesting.[46] In return for this lie, some were offered $1, while others were offered $20. While all of the subjects agreed to tell the lie, it was noted that those subjects who were offered the smaller sum of money ($1) rated the appeal of the test much higher in

[43] Stuart Sutherland (1992), pp. 232. Sutherland accordingly offers good advice for anyone considering writing a book: make your opening passages as good as possible, for (few books are actually read cover to cover, and) people will remember the beginning more than other parts, and will be influenced by it for the rest of their reading.
[44] Ibid, p. 25.
[45] Ibid, p. 26.
[46] Ibid, p. 102.

comparison with those who were offered the larger sum ($20). The explanation for this phenomenon given by psychologists is that a $20 sum is a meaningful bribe, whereas $1 is clearly not. Consequently, in order to self-justify their decision to mislead future participants, the subjects who were offered the lower sum had to actually convince themselves that the task they partook was *in fact* more interesting than they otherwise found it to be.[47] In contrast, those offered the larger sum were the recipients of a meaningful bribe, and accordingly they did not find themselves under pressure to alter their own point of view regarding the task in order to justify their deceit. Subsequent formulations of the experiment had the person being falsely informed of the interesting nature of the study refuse to partake in it. When the participants were asked after this how interesting they found the test, both those in the $1 and $20 groups now rated it at equivalent levels of dullness. We can explain this by noting that having not caused any harmful effect by influencing people into actions under false pretences, the $1 participants no longer required of themselves to believe the test was more interesting than it was, and so they were free to revert to their unbiased position.[48]

Other studies have found that 95% of British drivers think themselves to be better than average in their driving ability,[49] while 94% of university lecturers believe themselves to be better at their jobs than their peers.[50] Clearly these estimations cannot be accurate. Self-serving tendencies can be found manifest in the fact that 85% of medical students think it is improper for politicians to accept gifts from lobbyists, while only 46% of them think it is improper for physicians to accept gifts from drug companies.[51]

We tend to associate our successes with things and features largely intrinsic to us, such as the time we have spent training and studying or the personal qualities we have, but associate our failures with situational factors, such as untested new equipment or being tired in the exam.[52] These are of course self-serving determinations, as while our failures are mitigated by factors which we did not have strict control over our successes are placed firmly as belonging to us. We will also alter our judgments according to our decisions in order to validate them. So for example our estimation of the value of a house may increase simply because we have purchased it. This can be for several reasons, such as to negate a 'buyer's remorse', the tendency of people to regret a purchase immediately after it, and to help justify the decision to acquire it. The same thing can be seen in the acceptance of a new job. Soon-to-be graduates of a management course were asked in a study to evaluate firms which they had visited, and indicate their three most attractive job opportunities.[53] There was at this stage almost no difference between their indicated top

[47] Ibid, p. 103.
[48] Interestingly, other studies have shown that rewards for positive tasks can devalue them. For instance, in a study of 1200 adults asked to donate blood to a bloodmobile, many more *not* offered $10 to contribute agreed to do so than those who were offered $10 (Ibid, p. 106).
[49] Ibid, p. 240
[50] Thomas Gilovich (2001).
[51] Robert T. Caroll (2005), citing Ashley Wazana, 'Physicians and the Pharmaceutical Industry: Is a Gift Ever Just a Gift', *The Journal of the American Medical Association*, vol. 283 (2000).
[52] Stuart Sutherland (1992), p. 127. There is also evidence to suggest that we do the opposite with other people.
[53] Ibid, p. 93.

three. However, once the students had decided which job they were going to take, but before work had started, they were again asked to rate the firm's attractiveness. Not surprisingly, the job that they had committed themselves to was rated significantly higher after that decision than before, while the jobs which they had not elected to take were rated as less attractive.[54] Similarly, 'sunk cost error' describes a reluctance to pull out from commitments which we have already made.[55] For instance, if we were to pay to go and see a movie which turns out to be terrible, we may not want to leave halfway through even though if we had known that the movie was going to be that bad we might have paid a small sum of money to avoid seeing it at all. This can also account for a reluctance in people to avoid selling poor performing stocks, even where it is obvious that the stocks are likely only to lose more of their value. The significance of this tendency is substantial. Broadly described, an "inability to acknowledge our own errors even to ourselves is one of the most fundamental causes of irrationality".[56]

2.11 Wishful Thinking, Just World Phenomenon, and the 'Best of All Possible Worlds'

Voltaire's classic work 'Candide' mocked Leibniz relentlessly for arguing (via his metaphysical constructions) that this world is the best of all. His withering ridicule emphasised the counter-factuality of this world as being the 'best of all possible'. Yet like Leibniz we often do something of a very similar nature, albeit without the need for lengthy metaphysical treatises. Wishful thinking is a tendency for us to form pleasing beliefs, often without sufficient evidence and even contrary to the available evidence. Allied forces in world war two conducted an airborne operation of an area around Arnhem in order to capture a bridge across the river Rhine. Headed by the British General Montgomery, the plan was flawed for a number of reasons.[57] The strategy should have called for the capture of Antwerp first, and because it did not the Germans were able to move their 23rd Army into Northern Holland to take part in the defence of Arnhem. The strategy called for the XXX Corps to support the Airborne troops, however this task required them to move down a very narrow and restricted road about one tank wide (later nick-named 'Hell's Highway'), to either side of which was land that tanks could not operate in. Most significantly and interestingly for our analysis here, Montgomery chose to dismiss as 'ridiculous' reports of two (previously missing) German Panzer divisions being seen very close to the drop zone of the airborne troops. His staff followed him in this putatively military determination. Another, General Browning, commander of the 1st Army, upon being shown aerial photographs of German tanks in the area, remarked "I wouldn't trouble myself about those… they're probably not serviceable". Browning had doctors declare the intelligence officer who had brought the photos to his attention 'overtired' and placed on rest. The suggestion Sutherland makes here is that Montgomery, subject to all the same kinds of human tendencies as us all, and himself desiring to be able to claim being first over the river Rhine, was able to form his beliefs -

[54] Ibid.
[55] Note also the aforementioned 'boomerang effect'.
[56] Ibid, p. 100.
[57] Stuart Sutherland (1992), pp. 144-146.

contrary to much military wisdom – in ways that preserved the hope of the possibility of this.

Closely related to our tendency to form our beliefs in order to accommodate for ourselves a better world is the belief that the world is an intrinsically just one. 'Just-world hypothesis' describes a strong predisposition for us to believe that the world is an orderly, predictable, and essentially just place. The result of this can be judgments drawn that those who suffer are in some way personally responsible for that suffering. This belief can have ethically devastating effects, for as a result in some cases juries may feel that the victim of a crime was responsible for its occurrence. This was a likely factor in the case of a 22 year old Fort Lauderdale woman kidnapped from a restaurant and raped twice at knifepoint.[58] The jury was told that on the night the victim was wearing a short lace miniskirt, a tight top, and no underwear. The jury acquitted the man, finding that she was 'bound to attract attention' and had 'asked for it' by her choice of apparel. In a study a situation was set up in which a woman was seen by subjects being asked word association questions. Her failure to answer correctly would lead to her being electrically shocked as punishment. She was seen to continually answer incorrectly, and thus seen to repeatedly sustain electrical shocks. Observing the woman's plight, participants formed a low opinion of her, and even rated her as less attractive as a result of this as compared with before observing the testing.[59] It is likely to be simply psychologically uncomfortable for us to see others being harmed unjustly, randomly, or purposelessly, and as a result easier for us to deal with by ascribing some guilt or culpability on the parties' behalf.

2.12 Availability Error

Availability error exerts a formidable effect on human reasoning. The medical profession provides an illustrative example. Everyone would agree that doctors are highly intelligent people who possess extensive medical knowledge in regards to matters of health. All of them, we would think, are informed of the harmful effects of smoking to a person's health and the statistics relating to this. Despite this, smoking rates in general practitioners, specialists and consultants is higher than in those doctors who deal with the effects of smoking first hand, such as chest surgeons and radiologists.[60] What explains this fact? It is suggested that those in the latter group deal with the effects of smoking in far more salient ways, which in turn makes the consequences of smoking more cognitively 'available' to them than others. The visible consequences of smoking impress themselves more powerfully upon those doctors than others in the medical field, and so influence judgments relating to smoking more significantly, which in turn causes a greater aversion to smoking.

In a more mundane study, one group was asked which of the following was more similar: West and East Germany, or Sri Lanka and Nepal. This group answered that West and

[58] Claire Andre and Manuel Velasquez (1990).
[59] Ibid.
[60] Stuart Sutherland (1992), p. 24.

East Germany were more similar to each other than were Sri Lanka and Nepal. Another group was asked which was more dissimilar: West and East Germany, or Sri Lanka and Nepal. They answered in turn that West and East Germany were more dissimilar than Sri Lanka and Nepal. This creates an apparent contradiction of West and East Germany being both more similar and more dissimilar to one another than are Sri Lanka and Nepal.[61] The accepted explanation for this is that their relative abundance of knowledge in regards to West and East Germany in comparison with Sri Lanka and Nepal biases them in favour of answering West and East Germany in relation to both lines of questioning. Of American drivers who were told that they had a 0.0000025 chance of people dying in a single car trip, only 10% said they would wear their seatbelt. However, 39% of those who were told they had a 0.01 chance of dying *annually* in a car accident (a probability based on the same figure) said they would wear their seatbelt.[62] The notion of a 0.01 chance to die yearly is more available to us cognitively than the notion of a 0.0000025 chance to die per trip is, and so it registers a greater effect on our judgments when involved. Thus we can see that by choosing or controlling how information is represented to the public that judgments can be manipulated in very basic ways.

In a different study a group of people was asked to memorise the words 'adventurous', 'self confident', 'independent', 'persistent', while another was asked to memorise the words 'reckless', 'conceited', 'aloof', 'stubborn'. Both groups were then asked to read a short story about a young man who possesses several dangerous hobbies, is well abled, has few friends, and who rarely changes his mind. When asked to evaluate the man, and despite the fact that it was made clear to all participants that the words had nothing to do with the story, their judgments were seen to be nonetheless coloured in ways relating to the meanings of the terms memorized. That is, those who had memorised the more positive words came to see the man in a more positive light, and visa versa. The words were 'available' to them in their mental goings-on, and so subsequently influenced their judgments in strictly irrational ways.[63]

Another study saw subjects given 25 suicide notes and asked to determine which among them were genuine and which were fake. After stating whether they thought a note was genuine or not they were told whether they were correct. In truth none of the suicide notes were real, and had been made up for the purposes of the experiment. Half of the people were told that they were doing extremely well, while the other half were told that they were doing poorly.[64] After completion of the testing phase the nature of the experiment was revealed to the subjects, who were even shown the table by which they were randomly allotted to groups (i.e. whether they would be told they were doing well or doing poorly). When asked in a subsequent questionnaire how well they thought they would perform in the future if given real suicide notes, those who had been given the positive assurances of their accuracy estimated themselves much more confident in their ability to discern real notes from fake ones than those who had been given negative

[61] Thomas Gilovich (1991), pp. 36-37.
[62] Stuart Sutherland (1992), p. 227.
[63] Ibid, p. 17.
[64] Ibid, p. 148.

responses.[65] This in spite of the fact that each person knew that their assurances were predetermined. A variant of this experiment suggests that this effect was not caused by emotional personal attachment. The study had it so that subjects would watch *others* performing the same task of suicide note differentiation, and just as before half were told they were doing well, while the other half was told that they were doing poorly. With the test subjects watching, those who were giving the answers were - just as before - told that it was all a hoax, and that none of the suicide notes were real. Yet the observers still concluded that those who were given positive evaluations during the testing phase were more likely to be able to perform accurately with real suicide notes than those given negative evaluations. Significantly, if before estimating how well they would do in the future there was a group discussion of the ways in which beliefs tend to persist despite evidence to the contrary, it was found that this discussion had a marked effect on normalising the participants anticipations of future performances (although, it did not completely eradicate the phenomenon).

2.13 The Halo Effect

In 1969 Jerzy Kosinski's novel 'Steps' won the American national book award for fiction. Eight years later it was retyped into manuscript form by a roguish character and submitted to 14 major publishers and 13 literary agents (including the original publisher Random House) without a title and with a changed author's name.[66] Of the then 27 people who had read the manuscript not a single one recognised it as the original award winning novel. More significantly, of the 27, not a single one accepted it to be published. Explanation for this fact is supplied by the hypothesis that in the first case Kosinski's reputation coloured perceptions of the book in positive ways, and imparted onto it judgments that were not generated by the book alone. This effect is known as halo bias.

The stringent peer-reviewed academic journals can also be seen to come under the influence of halo bias. In a similarly themed story, in 1982 two psychologists took 12 well known psychology journals and selected an article written by someone from one of the top ten psychology departments of universities in the U.S.[67] As a result the articles selected were mostly written by eminent psychologists. Keeping the basic substance of the paper unchanged they altered the name and any other information within the text which would give away the true identity of the writer, put on them the name of an invented institution, and submitted the papers to the original journal of publication. After submission three journals recognised that they had already printed the papers. Of the nine remaining journals, eight rejected the submitted papers (all of which had previously been published – only under the original names of their eminent authors). The 8 editors and 24 referees each found that the paper they examined did not, in their view, merit publication. The conclusion drawn from this is straightforward: the name of the author and their institution can have a significant bearing on the light in which an academic paper comes to be seen by peers, so much so that if the name of the author and institution carries no

[65] Ibid, p. 149.
[66] Ibid, pp. 32-33.
[67] Ibid, pp. 30-31.

recognition then the work can be seen as unworthy of publication. In response to criticisms of this rouse, Sutherland remains adamant: it is clear that "the editor had either made a mistake by agreeing to publish originally or they made a mistake by not agreeing to publish subsequently".[68] There is a silver lining to be extracted from all of this: our failure in our pursuits may perhaps have less to do with the quality of our work and more to do with other factors. This however is dual edged, as it must also be admitted that our successes may likewise not be strictly merited.

A further study conducted by David Rosenhan is alleged to show the way in which professional expertise can be veridically undermined by halo bias. Eight sane study participants were sent to twelve mental hospitals in five different US states where they reported having heard auditory hallucinations. They told their examiners that the words 'thud', 'empty', and 'hollow', as well as other unidentifiable words were heard by them.[69] Seven of the eight were diagnosed with schizophrenia, and later discharged with 'schizophrenia in remission'. The average stay in the hospital was 19 days (they acted normally upon admission in order to be released as quickly as possible). The eighth among them was diagnosed with another psychological disorder. Notable was that patients within the hospitals were commonly able to identify the pseudopatients as faking their symptoms with a far greater success rate than the professionals were. Rosenhan suggested from his study that a halo effect had a role in explaining the inertia in not recognising the pseudopatient's sanity after admittance, as once labeled 'insane' it proved to be a hard classification to break.

Another hospital, upon hearing the results, boasted that their own would not make the same mistakes. Rosenhan therefore arranged so that the hospital in question was informed that within a 3 month period one or more pseudopatients would be sent to them to test if their claim was true. Staff at the hospital were asked to identify patients they suspected as being pseudopatients, and to give them a numerical rating based upon the degree to which their suspicions were raised. The results were as follows: "Forty-one patients were alleged, with high confidence, to be pseudopatients by at least one member of the staff. Twenty-three were considered suspect by at least one psychiatrist. Nineteen were suspected by one psychiatrist and one other staff member."[70] Rosenhan had in fact sent nobody to act as a pseudopatient, and none of those identified as pseudopatients had any connection to anything that had been advised.

2.14 Attractiveness Bias

In the 2004 Ukrainian Elections opposition candidate Victor Yushchenko claimed to have been poisoned by the Ukrainian government following a sudden disfigurement of his face during the campaign. It was later found that he was suffering from the effects of massive levels of dioxin in his body, a chemical agent known to be consistent with the symptoms he exhibited. In theory this could have been a smart ploy by anyone who had wished to

[68] Ibid, pp. 31-32.
[69] David L. Rosenhan (1973), pp. 250-258.
[70] Ibid.

see Yushchenko lose, for we do indeed possess certain biases in relation to the aesthetic qualities of others and which effect the way in which we think about them, and Yushchenko was known for his good looks. In practice however it would have been a mistake, as he used his apparent poisoning to rally public support behind him and ultimately win the election.

The role aesthetics plays in public life has been observed for some time. Studies have shown that attractive people have a social and economic advantage over others. Attractive people tend to be paid more, receive more attention, and be judged as more intelligent by others.[71] We find ourselves as tending to think that attractive people possess more positive personality traits, and expect them to do better in life in areas such as marriage and careers.[72] Attractive men and women are also more likely to be rated higher on their athletic prowess, sense of humour, and so on.[73]

Our preference for attractive people can be found reflected extensively in narrative. It is a cross cultural feature of folk tale traditions that young, reproductive-aged women are vastly overrepresented in stories involving female protagonists. This same pattern is repeated for males. Where described, female protagonists are almost universally described as physically attractive, as are males to a slightly lesser extent. Female antagonists are described as attractive less often (69% of the time, compared with 97% for female protagonists). Where there is a female antagonist, they are more likely to be older than 40 years of age. Additionally, female protagonists are regularly described as being less active in the pursuit of their goals as compared with male protagonists, and as very family-orientated.[74]

Attractiveness bias has potentially global consequences. In the case that a significant group of people are considered to be aesthetically impoverished according to the aesthetic standards of others, we may inquire as to what, if any, consequences this might have for judgments formed in relation to them on issues of ethics and social justice. I have heard it said with a noticeable uniformity by people that they consider Australian Aborigines to be, on the whole, the 'ugliest' race of people in the world. Now, suppose that these sentiments accurately reflect the aesthetic judgments of a large number of people. In such a case we may ask how judgments of this type against races and ethnicities effects their moral status and stems the empathy others have towards them. Such judgments are not likely, in my view, to be utterly benign in relation to questions of this type, and as such need address.

2.15 Conclusions

What this chapter has provided is a brief introduction into some of the serious and concerning aspects of human irrationality that result from cognitive biases. We have seen

[71] Greg Lester (2006).
[72] A.H. Eagly, et al. (1991), citing Dion (1972).
[73] Stuart Sutherland (1992), p. 28.
[74] Jonathan Gottschall (2005), pp. 85-103.

a number of ways from among many that human reason can fall into impassioned and unimpassioned error. I take it that, were it ever doubted, such considerations and findings as presented are enough to seriously upset naïve views of human cognition as rationally untroubled. As introduced earlier, one possible way towards a solution for the various problems raised by our irrationality could be to adopt an increased level of epistemic reliance upon expertise and authority, in order to guide us past our biases. So this argument goes, perhaps what we should seek to do is place greater responsibility in the hands of those whom we judge to be in a better position than we are to know – those who we determined to be more cognitively endowed and epistemically advantaged. This idea will be explored in the following chapter. It will be argued that this avenue does not offer a desirable resolution to the problems of irrationality outlined, and raises a number of independent epistemic problems of its own.

Problems of Expertise and Deference

3.1 Introduction

In 'Political Liberalism and Social Epistemology' Allen Buchanan uses his own experiences to illustrate the way in which our fundamental dependence upon others as sources of knowledge can place us at substantial moral and prudential risk.[75] Growing up in the heavily, systematically racist 1950s southern US, he recalls how family members and social figures inculcated him with various racist beliefs regarding those with black coloured skin. He recalls amongst other things having seen a severed black ear brought to school one day by another student, and joking with schoolmates about how black prison inmates had their genitals electrically shocked at a local penal farm. Escaping this environment mostly by luck at the age of 18, and then rejecting those racist attitudes as based upon false beliefs regarding natural racial differences between blacks and whites, he reports his first response as feeling bitterly betrayed by those in whom he had placed confidence and trust – parents, extended family members, teachers, pastors, and politicians. He later came to feel fortunate for having escaped this toxic environment before he had committed any serious acts of racism. What is essentially described in this example is the epistemic problem of 'vulnerability'– the situation in which the weight of responsibility we place (and must place) on others for a profound number of the beliefs we have about the world simultaneously allows for the possibility of our accepting false beliefs, some of which may place us at serious moral and prudential risk.

The problem of vulnerability is extended by the possibility of a society inculcating not only false beliefs, but strategies for overcoming cognitive dissonance caused by conflicting evidence. So for example if a society holds that black people are naturally inferior to white people, it may additionally instruct that instances of black people who appear at odds with this account (such as black people who possess some clear intelligence or learning) 'must have had white blood in them'.[76] In this way false beliefs can be systematically built into a person's way of thinking and given a rigid framework within which to survive recalcitrant evidence. With resilient ways of dealing with contrary evidence, racist beliefs and attitudes can develop to the point where social conditions positively support those initially held beliefs. If it is believed that black people are naturally intellectually inferior to others, they may be offered fewer jobs that would have provided opportunities to prove themselves to the contrary, or that would allow for the development of their intellectual abilities in more sophisticated ways.[77]

---

[75] Allen Buchanan (2004).

[76] Ibid, p. 96. Precisely this sort of effort in reasoning can be seen in the writings of David Hume. On the matter of a Jamaican man recognised for his 'learning' and his knowledge of 'parts', Hume chooses to sceptically dismisses the reports as most probably the result of exaggeration, thus maintaining his beliefs as is. See Eric Morton (2002).

[77] Allen Buchanan (2004), p. 96.

One way to work towards a reduction of false beliefs is to philosophically advance and endorse an optimal level of 'epistemic egalitarianism'. Epistemic egalitarianism refers to the "tendency of ordinary people to think well enough of themselves to be willing to challenge socially identified authorities on occasion, and to think well enough of their fellow citizens to be disposed to listen to them when they criticise socially recognised authorities and accepted practices".[78] This is contrasted with the condition in which society has "surplus epistemic deference", which is where it misplaces its deference to putative sources of authority, or has overconfidence in them in their ability to provide true beliefs. What is to be sought then is a balance in the inherent tension between, on one hand, the individual and social necessity for the division of epistemic labour and epistemic reliance, and on the other, the need for autonomous, epistemically scrutinizing individuals. Part of what can serve as a basis towards bringing about a greater level of scrutiny, or the capacity for it, is a wider social understanding of the historical, social, and epistemic characters of expertise.

There are two questions that we may ask in order to bring into sharper focus the questions before us: a) 'why should we doubt the experts?', and b) 'why should we trust the experts, and in what ways are we enabled to?' This chapter will explore reasons for which we should register an appropriate and grounded level of epistemic scrutiny in relation to expertise, while the third chapter will aim to outline answers to the second question.

It will be put forward that historically there are many good reasons to have concern for our dependence upon experts for beliefs, especially when of a morally or prudentially significant character. It will be argued that certain biases exist for epistemic authorities that can exacerbate the problem of our vulnerability, and in cases can do so precisely because of the nature of expertise and the social conditions that can surround it. Essentially, where the first chapter gave us reason to question our veracity of reasoning by cataloguing a number of significant truth-undermining biases and features of thought, the second will do the same in relation to our confidence in epistemic authority. With this outline in place a philosophical framework for projects that seek to limit surplus epistemic deference can be developed.

3.2 Problems of Intellectual Authority: An Historical Analysis

> It is not enough to have a good mind, the main thing is to use it well.
> --Rene Descartes, *Discourse on Method*.

In 'The Open Society and its Enemies' Karl Popper critically explores how the practice of deferral to the 'intellectual leaders of mankind' has opened up the good of society to substantial risks.[79] Wary of the history of philosophy, Popper wrote that "great men may make great mistakes… some of the greatest leaders of the past [have] supported the perennial attack on freedom and reason". Writing in a time of great conflict, Popper

[78] Ibid, p. 99.
[79] Karl Popper (1962).

stressed the importance of the issue, stating his belief that if "…civilisation is to survive, we must break with the habit of deference to great men". That the thoughts of great men went too regularly unchallenged, he argued, resulted in our continually being influenced by the mistaken views they advanced under the pretext of observing our "intellectual heritage".[80]

On the basis of his legacy Popper singles Plato out as being particularly guilty of infusing western thought with beliefs that put to risk our intellectual and moral good. While he openly admits that Plato was in possession of a profound intelligence, he contends that in him is also to be found a man held historically in such high regard that many thinkers since have been blind to a number of his greatest failings. Popper notes Plato's admiration for one of histories most brutal, repressive, totalitarian, authoritarian, and militarily powerful societies – Sparta – as being of particular concern and influence in the development of his views. Plato's ideal society would have been structured as a caste system where rulers ruled dutifully and where movement between the classes was restricted in order to maintain the purity of the 'race' of the master classes. Popper argues however that this would do little more than turn the greater proportion of the population into "human cattle", holding them in a slavish condition to the will of the guardians and the philosopher kings (who were putatively there to serve towards the interests of all). Plato encouraged eugenics that would have seen babies culled that did not meet the Hellenistic physical ideals of the time.[81] The ideal society would strictly control and limit people's exposure to literature and music to avoid their corruption, while individuals would be prohibited from the private ownership of property and precious metals in order that that possession would not lead to disunity. In line with this, women and children were to be considered the collective property of society. While women were to be educated just as men were, and allowed to participate in the ruling class, Popper noted that women were nonetheless seen by Plato as the degenerate reincarnation of "cowards" and "villains".[82] Plato advocated that the political class should be able to lie in order to suit its purposes, such as when deceit promoted the stability of rule.

This picture painted for us by Popper's understanding of Plato is something of an Orwellian nightmare, where the state interferes with and strictly controls the private sphere of society under the pretext of acting towards its benefit. But as much as Plato's influence upon the western philosophical tradition is argued to have been significant enough to make many of its most sinister failings difficult to see, Hegel's influence is outlined as being similarly detrimental. According to Popper, concealed within Hegel's monstrous use of language are endorsements for the subsumption of the individual to the state, for the use of ideology in accordance with the interests of the state, for historical prophecy, and for wars between classes (for the left) and nations (for the right).[83] Popper goes on to criticise Marxist philosophy, finding much there similarly disagreeable.

---

[80] Karl Popper (1962), vol. 1.

[81] Which ironically could well have seen his mentor Socrates killed at birth, given that Socrates was reported to have possessed an unsettling degree of ugliness. See for instance his descriptions in Plato's 'Theaetetus' and Xenophon's 'Symposium'.

[82] Karl Popper (1962), vol. 1, pp. 38, 48, 49, 52, 53-54.

[83] Karl Popper (1962), vol. 2.

How are we to begin to explain these intellectual blind spots? Popper suggests that much of the fascination with these great figures of our intellectual history is due to a sort of spellbound enchantment we have with them, cultivated in part by the philosophers themselves. "For some reason, philosophers have kept around themselves, even in our day, something of the atmosphere of the magician… philosophy deal[s] with those mysteries with which religion deals, but not in a way that can be revealed to common people; it is considered too profound for that".[84] Hegel's philosophy claimed that it possessed the ability to reveal the most profound truths, while simultaneously masking many of these in the most difficult and impenetrable language ever composed. "It knows all about everything", and "who can be sure that the answer is not true", given the complexity of the language involved?[85] Similarly, in Plato's Allegory of the Cave it is the philosophical sage alone who sees the truth of the 'real' world outside, and who there alone has access to secret and arcane wisdoms. This sage then dutifully acts as mediatory between truth and society, which has a clear quasi-prophetic ring to it.

Popper warns us through these examples not to be taken in by "mystifying nonsense", by the "magic of high sounding words, and by the power of jargon". We are advised not to allow ourselves to become entrapped in awe for great thinkers and high reputations, for these men shown us just how easily "a clown may be a 'maker of history'",[86] and may be made an "object of a kind of worship".[87]

Expanding our historical analysis beyond Hegel and Plato, we can find falsehood in the beliefs of many venerated thinkers. We can find that Aristotle believed males to be conceived in strong northern winds,[88] and that he believed women to have fewer teeth than men. As Bertrand Russell noted on that point, Aristotle was married twice but seemingly never bothered to check. Russell's own writings however reveal that he endorsed wars of colonial genocide under the appropriate conditions.[89] David Hume, a man of great historical and philosophical influence, was particularly disdainful and suspecting of the natural inferior quality of 'negroes'.[90] A young Voltaire was similarly sceptical of racial equality between Europeans and Africans.[91] Francis Bacon, who played an important role in early conceptions of science, believed that warts could be cured by being rubbed with bacon.[92] Arthur Schopenhauer curiously believed that reading excessively could ruin one's own ability to think independently, and holds the dubious

[84] Ibid, p. 33.
[85] Ibid, pp. 33-34.
[86] Referring here more to Hegel than Plato. 'Maker of history' has a dual meaning, referring both to Hegel's centre stage in the history of philosophy, and the philosophy of historicism and historical prophecy, of which Hegel was a famous proponent.
[87] Ibid, p. 36 (quoting with approval Schopenhauer on Hegel), 31, 36, 59.
[88] Thomas Gilovich (1991).
[89] Those being, namely, that the party engaging in genocide is 'culturally superior', and could put the land to better use. See Bertrand Russell (1915). The cause of pacifism, the case for which Russell is recognised as being a steadfast proponent for, is otherwise advanced in this paper.
[90] David Hume (1753).
[91] See his 'Philosophical Letters' (1733).
[92] Thomas Gilovich (1991).

distinction of having writing one of the most misogynistic essays of all time.[93] Rousseau believed that to fantasise about someone in a sexual context was an act of some meaningful vice, for it was in some way to rape their concept.[94] And a young Immanuel Kant insisted in a quite surreal display of racism that anything said by a person of black skin can be *a priori* known as stupid.[95] Later on he wrote that children born out of wedlock were effectively *untermensch*, sub-human parasites of the state that did not deserve nor should not receive the same moral status as other human beings, and who were to be treated as was seen fit by others. He believed further that masturbation was a greater moral infraction than rape, that women should prefer to kill themselves than to permit themselves to be raped, and that sex for anything other than procreation was immoral. In Kant was a genuinely outstanding intellectual, one of the singularly greatest thinkers of all time – a man revered for his capacity for critical thought and powerful intellectual insight, and yet who nonetheless held beliefs whose substance we might otherwise ascribe to people characterisable in terms of their ignorance, bigotry and thoughtlessness.

This short survey could surely be expanded upon greatly. It might be immediately objected that I am being unfair to these thinkers, and that I am ignoring issues of anachronism in criticizing them with the benefit of hindsight and modern view and values. Surely I cannot rightly expect thinkers of past ages to have in every way found the same moral and intellectual beliefs as are upheld in modern liberal societies. Of course this is true, but as an objection it misses the point and misconstrues the intention of the criticism. I would first simply note however that of all people whom this appeal might be made in defence of, the case of Kant is the weakest. Kant wrote the well known essay "What is Enlightenment?" saying that it was to think for oneself, to leave our self imposed intellectual immaturity, to 'dare to know' without need for the guidance of others. Furthermore, his whole moral philosophy was based around the autonomy of thought and the autonomous rational exercise of judgment. It is arguable that if we are to say that Kant could not think for himself on these issues (in order to rescue him from any personal injury he might sustain in our eyes by having direct ownership of his beliefs), then virtually none of us can be said to be the true, independent authors of our own thoughts, or held responsible for them, given the place independent thinking held within his mode of thought.

What these examples *prima facie* point us towards is no more than the relatively uncontested point that an unbounded deferential trust in the most exalted figures of our intellectual history on every issue would have been, both then and now, an ethically and veridically compromising dependence. This point is seemingly initially diminished by the fact that few epistemic authorities would express (publicly, at least) that they consider themselves or their peers the rightful recipients of such a wide ranging trust as that. We may reasonably ask then how it is that such failings are to be of any interest or value to us.

[93] Arthur Schopenhauer (2004).
[94] Alan Soble (2003), pp 55-89.
[95] Immanuel Kant (1965), p. 113.

My reply to that is this. Knowledge of cases in which outstanding intellectual figures came into both false and dangerously false belief instructively reminds us by example of our own limitations and room for error in thinking. By this knowledge we are enabled to more easily see how falsehood does not make allowances for reputation, grand intellect, or historical standing, and how being able to find truth in some areas assures sometimes very little in relation to being able to find it consistently in others. But to turn to the deeper issue here, while it may certainly be true that few intellectuals - if any - make explicit endorsements of their authoritative status on matters of some consequence outside of the specific domain of their authority,[96] that does not rule out their implicit encouragement or cultivation of that expanded deferential status in other ways, or even in their being afforded it completely innocently. The 'intellectual leaders of mankind' are precisely that, and through this social recognition they may become the recipients of a broad deferential status fully absent of any explicit appeals or encouragement of it on their behalf. Society, through recognising these people as being more able than others to come to sophisticated and developed understandings of the world, may recognise them on a similar footing in relation to their ability to generate, for example, true moral beliefs or to locate and exhibit desirable values. The distinctions between types of expertise and types of propositions may not be clearly held within any given society, and in that confusion there can be substantial room for the conflation of one expertise with another.[97] This may be further consciously or unconsciously cultivated by the way in which thinkers approach their audience. One powerful benefit then of a wider social awareness of the faults of intellectuals historically is to aid in the reduction of an overextended public trust towards epistemic authorities by providing resolute, memorable, and sometimes alarming examples of mistaken historical thinking amongst those who are and have been most admired. Lastly, it is a notable evolutionary strategy for us to seek to emulate the socially successful or prestigious,[98] but there is little reason why this emulation cannot extend towards the wholesale adoption of others' beliefs and values more holistically. Where people may be swayed towards modifying their moral beliefs and values in order to adopt those of notable individuals, and those modifications result in the uptake of false or pernicious moral beliefs and values, the epistemic situation clearly possesses harmful consequences. Taken together then these reasons issue grounds for us to seek to guard ourselves against being uncritically seduced by great minds and figures.

For those who are the targets of our deference, I periodically use the term 'Epistemically Significant Persons' (or 'People'; ESP's for short). They are those individuals or groups in whom epistemic trust is placed to provide true beliefs and/or right values, and through this trust-placing are in some way a significant epistemic influence. There are many

[96] I recognise of course that a theologian or a philosopher for instance is more likely to claim a deferential moral authority than is, say, a structural engineer, architect, or zoologist.

[97] Buchanan notes the distinction between status and merit based trust. Status trust is trust afforded to a person or group through their relationship to a recognised category of expertise (e.g. a philosopher). Merit based trust is conferred to an individual through appraisal on a performance basis. The two do not appear exclusive of one another, however, and it is possible for a person to gather trust in both ways. So, a particularly outstanding moral philosopher may be afforded a heightened deferential status relative to others of his kind of expertise for his individual performance, and yet simultaneously enjoy status and trust bestowed on the merits of his type of expertise. See his (2004).

[98] Peter J. Richerson and Robert Boyd (2005), chpt 5.

forms of expertise in which trust can be placed in this way, such as in religious, ideological, legal, parental, ethical, political, aesthetic, medical, or practical ways (to name several).

It is important to recognise that whoever may be an ESP for one individual or group need not be similarly so for others. For most New Zealanders the prophet Mohammad does not provide much of an epistemic influence at all, in terms of acting as a source of metaphysical, ethical, and social beliefs. Yet for billions of people elsewhere he is of profound significance in this respect. Indeed it is precisely by knowing what Mohammed did, said, and approved of (as recorded in the Hadith) that many fundamental beliefs about ourselves and the world are acquired and supported by Muslims. Within broad social groupings, we can readily imagine that while Martin Luther King bears a fairly extended epistemic influence on Americans generally, he also generally provides a much more potent epistemic influence on African Americans specifically, for psychological and historical reasons.[99] Epistemic relationships then, as these few examples illustrate, are dynamic, and vary according to the contingent particulars of the social groupings which we associate ourselves with.

3.3 Problems of Intellectual Authority: A Psychological and Theoretical Analysis

Karl Popper, noted before as having been deeply concerned about the problems surrounding intellectual leadership, contributed in significant ways to philosophical discussions of the nature of scientific theory.[100] His theory concerning scientific discovery painted it as principally engaged in by seekers of the truth who would go about advancing tentative hypotheses, finding corraborative evidence, and then passionately and boldly subjecting their theories to rigourous testing phases in order to falsify them.[101] Scientists as individuals were recognised as subject to all the same sorts of biases and prejudices as all people are, but through the distinct social dimension and engagement of science, and the subjection of scientific theories to on-going communal scrutiny, it was raised above these individual features to become in its own way an objective practice. Popper thus stressed the importance of science's intersubjective character as crucial to its having an objective standing.[102] The use of logically falsifiable propositions, and the act of seeking the falsification of hypotheses and theories, were two key components in what it was in Popper's eyes to being doing science. The greater the empirical content of the theory, the more opportunities it opened up for it to be shown to be false, and so the more scientifically valuable it was in that way.

---

[99] Such as that it is easier to identify with him in the case of shared racial identities, and the obvious fact that he worked specifically in order to face systematic injustices that black American's were subject to.

[100] Indeed, much of the basis of his criticisms of historicism and holism proceeded from his thinking in relation to the philosophy of science.

[101] See, for instance, Karl Popper (1963b), for an allusion to his conception of science as centralised around falsification and the pursuit of it; and Thomas Kuhn (1970a), for an account of Popper's theory and criticism.

[102] See Popper (1962), vol. 2, chpt. 23, for a picture of Popper's view of science as essentially social.

Thomas Kuhn came to a number of quite different conclusions about science. Kuhn held that if we are to arrive at a well developed understanding of science then we must observe how science has proceeded historically, and have that historical character reflected in the theory of science itself. In light of this view Kuhn recast science as a practice very much involved with the psychological and social landscape of its practitioners. Against Popper, Kuhn described science as a discipline which for the greater part seeks to maintain itself according to its accepted understandings. Kuhn termed this kind of science 'normal science', which was that done over the extended period in which core scientific 'paradigms' were not seriously brought under challenge by scientists (which was most of the time, according to Kuhn). In such times the failure of scientists to, using his terminology, "solve puzzles" from within the context of that paradigm is considered by the scientific community to be the fault of the practitioner rather than of the underlying theory, thus isolating and shielding the core theory from refutation.[103] It can be put even more simply: it is not the *theory* but the *scientist* that is tested during times of normal science. Eventually however incongruities between a paradigmatic theory and observations that are made develop to a point in which the theory itself begins to be questioned. In these tumultuous times a period of rare 'revolutionary' (or 'extraordinary') science is entered into, and new proposals for a replacement paradigm are offered.

Kuhn claimed that scientific knowledge during times of extraordinary science is not incrementally cumulative. Famously (or perhaps notoriously), he argued that a paradigm shift is a largely non-rational thing, as it involves switching between theories that employ the use of incommensurable theoretical concepts and terms. According to Kuhn, in arguing for a paradigm the doctrines of that paradigm are (and must be) referenced, but as it is precisely these which come under question during times of revolutionary science, the argumentation is of a partly circular nature. This in turn necessitated that theory choice is resolved in the end not by shared rational principles, but by psychological and sociological ones.[104] This controversial picture of science then sees it as something deeply rooted in the subjective aspects of those who engage in it.

Imre Lakatos was in some way critical of both the Popperian and Kuhnian accounts, holding that "the history of science refutes both Popper and Kuhn".[105] Lakatos however, like Kuhn, also iterated the self-preserving character of scientific theories in relation to the psychological characters and theoretical commitments of scientists. He held that the core of a paradigmatic scientific theory can be found protected by a belt of more malleable periphery theory which serves as a buffer to protect it from falsification by inconsistent observational data.[106] A further positive heuristic exists which tells scientists both how to do science (to 'problem solve') in light of the core paradigm, and also how to modify the protective theory in light of recalcitrant evidence. Lakatos noted that as a rule

[103] Thomas Kuhn (1970a).
[104] See Thomas Kuhn (1970b). Later on Kuhn indicated that there were indeed "transparadigm criteria" of an objective character that can serve as shared standards in theory choice. Kuhn reports being "surprised" that he was interpreted as suggesting otherwise. These criteria though, Kuhn held, are insufficient to fully determine the case, and still involve in their application subjective elements. See Thomas Kuhn (1977).
[105] Notes added to a transcript of a talk later published. See Imre Lakatos, 'Science and Pseudoscience (transcript)', retrieved http://www.lse.ac.uk/collections/lakatos/scienceAndPseudoscienceTranscript.htm.
[106] Imre Lakatos (1977).

scientists only ever seem to talk about anomalies to a theory, *not* refutations. Scientists, he said, can always seek to add further *ad hoc* theory to a research program in order to avoid a refutation of the more fundamental core theory by observation, thus aiding in the preservation of the underlying theory that they bear a commitment to. Without any need to endorse a particular picture of science such as those presented above, it is agreed with that the practice of science is meaningfully influenced by various psychological and sociological factors. This has implications for discussions of epistemic dependence.

Much of the discussion to come involves itself with science. The intention with this is not to draw out science as being particularly in need of attention, or to mark it as especially fault-ridden. Science, as perhaps the most venerated of all forms of inquiry available to us, allows us to see that even our most powerful ways of learning about the world can play host to a number of significant epistemic concerns. We might well think that even if scientific inquiry is not immune from the influence of biases, that at least it is in a better position than other kinds of inquiry, since there are various institutional arrangements in place within the scientific community to subject findings to criticism and to engage in a critical process *precisely in order that* bias and error can be revealed as such. This is indeed something to greatly appreciate about science as practiced, and very likely part of the reason as to why it is so effective at providing us with true empirical beliefs. The salient point to be carried away from this though is that the biases which can be found within science nonetheless constitute a significant consideration within the framework of determining the best course of action to take in regards to our shared epistemic imperfections and vulnerable epistemic relationships. It is notable that even in light of such epistemic rigors, and attempts to minimize or stamp out the effects of bias, that its presence is still visible. From the specific examples given, inferences can be drawn regarding the epistemic possibilities inhering in expertise generally, and the significance such biases hold for organising and regulating our epistemic dependence.

First among the many problems of expertise is the way in which self-serving motivations can influence inquiry and findings. The existence of some degree of self-serving motivation on the part of truth-inquirers is unsurprising, but recent psychological research has argued for a quite deep connection between scientific pursuit and the various social benefits on offer, and thus attempts to establish that self-serving interests are quite central to motivations for engaging in science at all. Satoshi Kanazawa analysed the biographies of hundreds of scientists, musicians, painters, and writers, and found that a number of common trends can be identified between them.[107] In the case of 280 randomly sampled scientists (98% of whom were male), productivity was shown to decline rapidly with age. The majority of them (65%) made their most significant contributions to science before their mid-thirties, while 80% had done so before their early forties.[108] The study found that peak scientific productivity generally occurs within a temporal window of 12 years surrounding the age of 30. This same pattern was found in the case of male jazz musicians, modern painters, authors, but also interestingly in the case of criminals. Kanazawa argues that it is the same mechanism that produces the "age-genius" curves that is responsible for the invariant "age-crime" curve. In other words, "both crime and

[107] Satoshi Kanazawa (2003), pp. 257–272.
[108] Ibid, p. 259.

genius are expressions of young men's proximate competitive desires".[109] Pursuing success in either scientific research or criminal activity are both ways towards acquiring the kinds of social rewards and resources that aid in attracting members of the opposite sex (genius would appear more focused upon acquiring status, while criminality in increasing available resources, although not exclusively). Contrary to expressions of genius, however, criminality requires no unique talent, and thus is principally available to a wider number of people as a means towards competitive genetic ends. The connection between social status and one's ability to attract mates is well known, both within human societies and lower order social animal societies. Kanazawa's argument is supported by the fact that marriage appears to depress the 'productivity' of both criminals and geniuses within roughly 5 years. Marriage has in fact long been known to be one of the strongest predictors of a desistance in criminal activity,[110] and Kanazawa's research shows that scientists who never marry do not appear to suffer from productivity depression.[111] If this picture of the motivations of truth-inquirers is accurate, then the account which has scientists engaging in research at least partly in order to obtain socially available rewards is significant.

That self-serving inclinations do play a significant role in motivating scientific research need not be a depressing realisation or admission, for this facet of truth-inquiry can be seen to have notably beneficial effects for the epistemic community in some ways. Philip Kitcher has argued that self-serving interests might lead scientists to explore alternative lines of research outside of those which have wide acceptance within the scientific community in order to obtain for themselves the social goods possible in pioneering new and promising research.[112] This natural distribution of effort may happen benignly, in an unstructured fashion. A new-coming scientist may survey a domain of interest in order to ascertain which areas of research are over competed for, and take up research in any under competed areas they identify. While this behaviour may be primarily motivated by self interest, it is also likely that for scientists to uniformly commit to a singular line of research into a problem is a wasteful mode of social truth inquiry. By not clumping together in research homogeneity is avoided, and the resulting diversity of approaches likely closer to what the optimal division of epistemic labour would be. Kitcher however warns against drawing global conclusions about the *on-balance* (i.e. positive or negative) role competitiveness might play in truth inquiry.[113]

The division of epistemic labour need not always proceed simply from elementary self-interested motivations. It could easily be that a society could proactively ask a group of scientists to follow out research into theories (and, Kitcher oddly suggests, even *believe* in those theories) that are not accepted by the wider scientific community, given the potential benefits obtained in non-conformity in truth inquiry. The benefits of this kind of division of scientific labour may be most salient during times of war, where successes in

[109] Ibid, p. 265.
[110] Ibid, p. 269, citing Laub, Nagin, & Sampson (1998), Sampson & Laub, (1993) (albeit, with alternate explanations to the one Kanazawa supports).
[111] Ibid, pp. 268-269.
[112] Philip Kitcher (2001), chpt. 9, and Kitcher (1990).
[113] Philip Kitcher (2002).

research can be of the utmost critical importance to collective and individual welfare. The history of science records such occasions, such as during the Second World War when requests were made of scientists by governments for research into diverse areas with potential benefits for the war effort. Paul Feyerabend argued a similar point regarding the value to society of theory diversity, arguing that divergent or dissenting theories to those most widely upheld should not be dismissed out of hand by the scientific community, given the value of heterogeneity in truth inquiry.[114]

Despite the positive epistemic effects which self-serving psychological principles can offer, their expression can threaten to degrade the veritistic connection of an epistemic authority. To the extent that trusted expertise yields social rewards there is an interest in accentuating and expanding the boundaries of one's expertise. This can lead to experts making claims towards expertise or claims that implicitly coattail upon their expertise outside of their epistemically appropriate area. Moreover, where their expertise is difficult to come by, such as where it is in a highly technical or difficult area of specialty, their efforts in obtaining that expertise can slew their judgments regarding its worth and veridical capacities, thus exaggerating its usefulness.[115] As mentioned in the previous chapter, this kind of psychological effect is well known: if we buy a new house or accept a new job, by that very act we will then likely find ourselves afterwards finding many more things pleasing to us about the house or the job than before we acquired them. It is part of our need to retrospectively justify our decisions to ourselves, which in turn can distort the judgments we make. Likewise, any particular scientific theory will probably have a great deal of time and energy spent in developing it, seeking corroborative evidence, developing supportive lines of argumentation, and thereafter in dealing with recalcitrant evidence and external criticism. Outside of this energy investment is the potential that the success of the theory can have for an individuals' sense of self-esteem, social standing, and financial position. This kind of natural psychological involvement and investment in the truth of a theory or domain of expertise can provide the fertile soil from which factors can arise which work against sound or balanced thinking on the matter.

One of the well appreciated consequences of the theory-ladeness of observation is that it allows for a particular theory to be retained in the face of recalcitrant evidence. This is possible through the fact that individual scientific theories are testable only in conjunction with auxiliary theory. Newton's theory of universal gravitation by itself, for example, tells us nothing about the locations of the celestial bodies of our solar system. For that, one needs to bring in the three laws of motion, the relevant initial conditions (such as the mass of the sun, the position and velocity of a particular planet or object of interest, that body's distance from the sun at some particular point in time), and make a number of key assumptions, such as that the body is not being affected in any serious way gravitationally or kinetically at the time of observation. To actually make

[114] He applied this sort of reasoning not only to theory content but also to methodology and theory selection. Feyerabend stressed that the flaws of some particular theory may only be discernible from the different theoretical stand-point (and even, from a *false* one). It was from these sorts of considerations that his philosophical motto followed, 'anything goes'. See Feyerabend (1978).
[115] Allen Buchanan (2004), p. 104.

observations one needs to utilise a number of complex and theory-dependent instruments, such as clocks, cameras, and telescopes.[116] As a logical entailment of this, a negative observational result implies nothing in and of itself about the falsehood of *some particular piece* of the total theory relied upon to make predictions and observations, rather than holistically the entire body of theory which was used. This epistemic feature is called the 'ambiguity of falsification'. On the basis of the theoretical networking necessary for scientific prediction and observation, it is logically possible for a scientist to avoid the refutation of his or her cherished theory by negative observation by issuing the point of falsehood at some ancillary point outside of their preferred theory. Whether this can be done indefinitely, and *reasonably so*, is another question, however.[117]

At a higher level of analysis competitive psychological factors can be seen again, this time at the level of the group. Alvin Goldman has noted the possibility that experts may tend towards encouraging a more positive judgment of themselves in others in relation to the value of their expertise.[118] Scientific groups in competition with one another for limited resources are potentially disposed towards presenting their findings in a more favourable light than can be justified objectively, in order to advance their funding prospects.[119] We can accordingly anticipate a number of ways, at both the individual and group level, that self-serving psychological traits can exaggerate the benefits, scope, and capability of an expertise, and thus the worth of the expertise as a whole. This in turn aids to cultivate an increased level of trust in the expertise, but in a way that only further adds to problem of excess epistemic surplus, for it occurs for reasons relating to the psychological features of the experts rather than the veritistic connection of the expertise.

In conjunction with certain social arrangements and views, self-serving tendencies may inculcate modes of thinking that are inhibitive of our capacity to be self-critical, as has been noted by Buchanan in the case of professional medicine.[120] Some doctors up until the mid 1970's in the U.S. expressed the view that they had special rights to withhold information from patients, or even lie to them, where they viewed such acts as being in the patients' best interests.

In order to justify medical paternalism arguments were advanced in support of it, but these were immediately found to be transparently in error by observers. Some for instance equated the medical good of a patient with their most fundamental good. But this is clearly a distorted view of what counts as good; our good as agents cannot be reduced to any single element, much less plausibly so that of our condition of health. We often quite reasonably make judgments and take actions that involve putting our health at increased risks (e.g. drinking alcohol, eating certain flavoursome but unhealthy foods, engaging in certain activities (e.g. sports, race car driving, sex, sitting in the sun), acting at our own risk to save a person in danger, and so on) in order to advance some other

[116] Martin Curd and J. A. Cover (1998), pp. 356-357. Any testing procedure that relies upon non-natural language as part of the observation (e.g. amp, volt, degree) can be immediately seen to be theory-laden.
[117] Duhem was at pains to point out that there were limits to the rationality of forever modifying one's theory in the face of recalcitrant evidence, whereas Quine seems to have thought that there are none.
[118] Alvin Goldman (2001).
[119] Ibid.
[120] Allen Buchanan (2002).

good or value we hold. Another chief argument held that when a person goes to a doctor they sign a theoretical contract in which they consent to surrendering their judgment to the doctor on all medical matters. But this is plainly no truer than it is the case that by taking our car to a mechanic for repairs we give them licence to make whatever alterations to our vehicle they see fit to, regardless of what our personal preferences may be, or what such alterations may cost us financially.

The fact that one of our most central and essential goods – our autonomy – was dismissed in order to make way for medical paternalism is worrying, and prompted concern from observers. But this in itself was not quite as worrying as the realisation of *who it was* that offered these flawed arguments. The deeper question provoked here was why "otherwise highly educated and intelligent individuals who routinely engaged in sophisticated chains of reasoning regarding scientific matters offer such transparently defective arguments to justify their paternalistic behavior".[121] From analysis it would be suggested that they exhibited a "selective cognitive disability", a reduced ability to be self-critical, and were accordingly all too willing to adopt self-serving beliefs in order to serve their own interests in matters.[122] They were morally and intellectually undermined by the sense of their own importance and their privileged place within society. They viewed themselves - and indeed were viewed by others, and so were externally reinforced in that view - as elites of society; intellectually, professionally, and economically on a higher (and better) social echelon than those who came to make use of their services. When it came to them believing something that would have theoretically done away with a patient's grounds to disagree with them on matters regarding medical treatment, that belief was able to pass through critical inspection without being flagged. It was philosophically and ethically significant that those who had advanced paternalistic arguments were found to be less than receptive to criticism over the issue, and tended to ignore philosophical challenges. But this seemed only to further compound their error. This response was similarly explained as resulting from that same 'higher-than-thou' mentality which had played a fundamental role in bringing about the issue in the first place, and which was now to a certain extent insulating them from external criticism. By this example we can see that social conditions can play in fundamental role in the production and support of false moral beliefs.

It is possible to spot an additional bias in effect here, as one may well suspect that self-deception had a role in the development of their rational and moral blind spot. That is, those in favour of such paternalistic measures may have more easily allowed themselves to believe on some level that their ethical judgments were genuinely intended to be in the greater interests of the patient, as opposed to their own. In that case it would be little different from the example of a "sadistic schoolmaster who believes that he beats small boys for their own good rather than to indulge his own erotic desires".[123] Were they able to observe with greater clarity how such an ethical position as the one endorsed catered so conveniently to their own interests, they may have been able to apply greater critical attention to the matter.

[121] Allen Buchanan (2002), p. 132.
[122] Ibid.
[123] Example borrowed from Sutherland (1992).

At a higher level of analysis we can find that the shared beliefs and understandings of a domain of expertise hold the potential to shape judgments in ways that threaten an expertise's veridical capacities. Theoretical paradigms, the foundational principles of a discipline of inquiry, are widely understood as being an important means for centralising and normalising concepts and understandings, and core to advancing knowledge by providing a relatively stable space from which to work and make progress. But just like the self-serving interests of experts they are dual-edged in that they offer the simultaneous possibility of inhibiting veridical content through confining understanding along limited dimensions. Scientists, academics, and professionals can be deeply invested in the central paradigm of their domain, both emotionally, psychological, and theoretically, such that through that dependence it can be difficult (or even theoretically impossible)[124] for them to examine things outside of it. The effects of this commitment will depend in a large way upon the specific nature of the paradigm in question. An alleged instance of the way in which such commitments can negatively influence truth inquiry at a theoretical level is described by an unnamed (supposedly for employment related reasons) and disgruntled anthropologist, who wrote to Steven Pinker regarding the state of modern anthropology. Heavily critical of how many anthropological accounts are simply false, he or she writes,

> ...their professional "expertise" has made them complete and total gulls. Just as fundamentalism disposes you to accept accounts of miracles, being of the trained anthropologist faith disposes you to believe in any exotic account from Elsewhere. In fact, a lot of this nonsense is part of the standard intellectual equipment of every educated social scientist, providing a permanent obstacle to balanced reasoning about various psychological and social phenomena.[125]

This phenomenon need not be taken as anything particularly unique to the case of anthropology, if accepted, but an example which instances a more general theoretical possibility inherent in the acceptance of a paradigm. That is, by way of simply coming into some centralised core of understanding that a body of inquiry holds as core to itself, we can to that extent be withheld from easily or clearly seeing outside of those principles governing conceptual organisation and interpretation.

Connected to this idea is one which suggests that a domain's theoretical principles are not things that stand alone in some socially removed space and decided by objective measures alone, but are to the contrary very much open to the influence of background political, social, and ideological factors. It alleges that how science is engaged in is meaningfully and potentially detrimentally subject to the particular socio-historical context from within which it arises and operates. The history of statistics provides a salient example of the way that ideological commitments can move the direction of truth inquiry in this way.[126]

---

[124] Refer to Kuhn and Feyerabend for various (controversial) claims regarding theory incommensurability.
[125] Steven Pinker (1994), p. 412.
[126] Alvin Goldman (1999), p. 39, who draws upon the work of Mackenzie.

Eugenicists in Britain in the early 20th century advocated social programs aimed at weeding out the socially 'undesirable'. They viewed 'undesirable' aspects of certain segments of society as being genetically inherited, and so encouraged measures to be taken in order to provide for the appropriate social conditions wherein the future influence of these genes could be limited. For example, they called for programs that promoted the fertility of "better types" while limiting the birth rate of the "unfit". Tax incentives and family allowances for the professional class were sought for their having children. Other proposals took a much more extreme direction, in the form of coercive negative eugenics.[127] Charles Darwin's cousin, Francis Galton, was one of those involved in the movement, and it has been argued that it was in direct connection to his social Darwinist beliefs that he worked to develop a set of concepts in statistics to deal with the variability of human characteristics in populations. If so, his example evidences a clear instance in which the background ideological positioning of truth inquirers can directly influence the direction and findings of science. As a feature of science this is something that most contemporary philosophers of science would not disagree with. Science, it is widely held, is best understood not as occurring within a social vacuum, but rather as intertwined with both the cognitive and contextual values of its practitioners. Some however have gone on to draw much stronger conclusions about the nature of science on the basis of its social connection, and attempted from examples such as Galton's to radically undermine confidence in science. Postmodernists have been particularly recognised for arguing in this way. It has been suggested by postmodern thinking that science cannot obtain truth, as all science is by its very nature biased in these sorts of background ways.

It simply does not follow, however, that because there were certain (in cases, unseemly) motivations for engaging in science research (or, in Galton's case, statistics with scientific application), that the findings of science are a construction of an ideological program which reflect nothing of the world. The fact that an individual or a collection of individuals can have special interests in the scientific research they engage in does not *a priori* entail that any or all of their findings run contrary to the truth of the matter. In order for this to be so a number of quite implausible assumptions need to be made, such as that biases always run contrary to truth, and that motivations alone determine what it is that people come to believe.[128] Rightly rejecting both of these assumptions, it is easily possible for us to see science as *both* contextually embedded *and* well able to arrive at empirical truth. Galton's example in fact illustrates this point very well, for the concepts he developed continue to this day to have wide ranging applications outside of what he originally intended them for.[129] It is worthwhile to note in that connection that few of those who use his formulas today are likely to know the historical details or ideological motivations that surrounded their initial inception, and more importantly that they *need not* possess that knowledge in order to use them successfully. While the strong claim here fails, the social character of science is an essential point worth retaining outside of these kinds of extreme employments. It enjoins on us an awareness of the way in which science

[127] See Allen Buchanan (2007), pp. 22-45, for discussion of the ethical and factual terrain involved in eugenics, and the connection between the two.
[128] Alvin Goldman (1999), p. 37.
[129] Ibid, p. 39, citing Alan Chalmers.

may be subsumed under the disparate background interests of those who engage in it. Because it is in no way necessary that bias is aligned with truth, we are advised to consider this in taking into account the broader epistemological issues involved.

A more modern example of this epistemic point comes the sociobiology debate, which vividly exposed the influence that politics and ideology can bring to bear in directing and socially regulating scientific research and academic debate, and raises deeply interesting and important philosophical questions about the extent to which truth may be pursued when it is perceived to come at a significant social cost. The so-called 'standard social sciences model' (the SSSM) was a fusion of ideas from anthropology and psychology, ideas which served according to Steven Pinker as "the secular ideology of our age".[130] The SSSM held that we are largely blank templates (or 'blank slates'), deriving much or most of our specific psychological/cognitive properties and content from the sociological and cultural conditions that surround us. Not only was this theory widely believed to be true, it was also thought rather important that it *was* true, for if it were then what had historically been taken as much of the justification of so many injustices (such as supposed natural, ingrained differences between sexes and races) could be rejected as simply false. Sociobiology, as introduced by E. O. Wilson in 1975, challenged this accepted view by bringing into focus a relatively benign yet markedly contrary claim: that evolutionary forces have acted to condition the behavioural tendencies and cognitive faculties of humans just as they have been widely accepted to have done in relation to our physiological features. According to this idea it is not just our contemporary social and physical environments which give rise to our 'humanness' or 'human nature', but the environments in which our ancestors evolved, and which have in turn bestowed upon us a genetic behavioural and cognitive endowment. As a field of study then sociobiology was in Wilson's words a "systematic study of the biological basis of all social behaviour".

Wilson's thesis set off a remarkable firestorm of criticism. For opponents his idea was anything but benign, and it drew significant exception from many; yet in so doing the statements and position of sociobiologists were construed by some in quite inexplicably false ways, drawing from their writings unjustifiably wild and vivid insinuations and implications. Kurzban surveyed the worst of the literature, and has relayed from there that "evolutionary principles imply genetic determinism", and that sociobiologists dismiss "cultural, historical, and individual variables".[131] Some have gone so far as to state that evolutionary psychology as a discipline of study is "a vision of morality and social order, a guide to moral behaviour and policy agendas", and even more obtusely, that it is "transparently part of the right-wing attack on collectivity" and the welfare state.[132]

For Kurzban that misdirection and intellectual ill-foundedness is as mystifying as the fact that it was and is academically tolerated, especially given that it endured in the face of the explicit denials of those who wrote the material that served to inspire it all. According to Kurzban, sociobiologists have continually "bent over backwards" to bring attention to the

[130] Steven Pinker (1994), p. 406.
[131] Robert Kurzban (2002), citing Nelkin (2000), Herrnstein Smith (2000), Rose (2000), and various others.
[132] Ibid.

fact that in their way of thinking cultural, social, and individual factors most assuredly play significant and meaningful roles in the development of human cognitive traits, and stress that in fact *nobody* thinks contrary to this.[133] So far as the sociobiologists are able to appreciate, the intellectual position that opponents have set their sights upon simply does not exist. Yet appeals to the fact that the kinds of claims alleged were never made, and that explicit statements were made to the contrary, have been ignored to the sociobiologist's ongoing frustration.[134] The costs of this are not insignificant, as it uses up time and resources, presents a false and disparaging picture of this avenue of scientific inquiry to the uninformed, and draws attention away from genuine points of disagreement in which debate could otherwise proceed.[135]

In order to understand the specific problem here, it is instructive to clarify that sociobiology could indeed be a discipline largely absent of truth. Let us assume this for the moment, in order to see the issue here in its clearest light. Even given this fact, this is something about sociobiology which stands quite separable and tangential to the issue of the way in which cases of argumentative attack proceeded. It would still be true, in other words, that while ultimately correct in relation to the truth-content of sociobiology, that in giving voice to their objection a number of critics engaged it with superficially poor argumentation in ways that, other things being equal, we normatively would have expected them not to. Knowing them beyond doubt to be highly intelligent and knowledgeable people well accustomed to thinking critically and reflectively, we would otherwise expect that such flaws would have been self-recognised before being brought forward. The fact that they were not points us towards the saliency of the issue of bias, and how it can contribute towards an abject failure to widely recognise and call attention to unsound argument and argumentative tactics.

If Kurzban's analysis is correct it draws further attention to the fact that the rightly venerated institutional measures of academia designed to encourage criticism and weed out bias - in this case at least - failed to do both equally well. If we understand those aims as distinct but related values, then it would seem that in this contentious example the social value of criticism outweighed the epistemic value of the identification and amelioration of the impact of bias on scientific and academic discourse. All of this in turn raises deeper philosophically important questions. Scientists engaging in research deemed by the majority to be politically or ideologically suspicious may incur highly significant social costs in doing so. In such cases, the pursuit of truth may take on a quite different dimension to that presented before, in which scientists engage in inquiry partly in order to obtain social rewards. For example, if a scientist believes that their contemplated research choice would be found by many peers and/or outsiders to be highly unpalatable (such as in the case of destructive animal research, or research into human racial differences), but anticipates greater goods available in completing that research successfully (say, the possibility of finding cures for significant human diseases, or in boldly discovering and facing the truth of matters, wherever it may rest), we may puzzle over how influential

[133] See Kurzban (2002) for a catalogue of the claims made against sociobiology and its modern derivatives, and replies offered by proponents of these theories.
[134] Ibid, p. 105.
[135] Ibid, p. 106.

such retributive factors will be in connection with the problems they select. Moreover, we may question the extent to which scientific research *ought* to be subject to legal, political, or social restrictions. When should the quest for truth take a back seat to concern for the social good?[136]

There are other ways in which social factors influence the direction of inquiry. What is taken to be of evidentiary significance depends in a strong way upon our background beliefs, which in turn must necessarily bear to some extent a social character and origin. In terms of a specific example, Kathleen Okruhlik argues that socially constructed gender preconceptions coloured the findings of the biological sciences for two centuries in relation to sperm-egg fertilisation.[137] She argues that certain social preconceptions regarding the gender differences of males and females, such as that males are active and females passive, blinded scientists to the possibility of the active role that eggs could play during conception. As early as 1795 the 'passive-egg / active-sperm' characterisation had appeared, mirroring the existing male-female social gender distinctions. In 1895 microvilli from sea urchin eggs extending towards sperm and attaching to them were discovered. However, this was largely ignored up until only relatively recently. Okruhlik takes this to show the more general point that theoretical preconceptions inform what questions we ask, and influence us with respect to what hypotheses we investigate and what evidence we recognise as being challenging to a theory. Equally relevant is that what criteria scientists take as constituting a good scientific theory will influence what theories are selected by them, but these criteria themselves will be open to contextual, social influence. Thomas Kuhn identified five core shared standards that scientists look for: accuracy, consistency, scope, simplicity, and fruitfulness.[138] But these standards, he argued, are open to change in terms of how they are ranked and applied by scientists in theory choice, and as such they are unable to be applied in theory selection in any logically strict or determinative way. As a result, Kuhn held, they are best thought of as *values* that are contributed towards by the particular social and historical situation in which they are applied. If further contextual values are added to the list of what constitutes a good theory (for instance, 'those theories are best which have either no negative social effect on the way people conceive of themselves or others, or a positive effect in that regard'), then those theories upheld by the scientific community as good can be even more clearly seen to be contextually influenced.

As a result we are guided away from conceptualising of science as in all times and places the unimpassioned, socially, historically, and politically disconnected pursuit of empirical truth. Yet it is epistemically significant that these underlying commitments, so far as they are consciously held, are seldom openly or forthrightly disclosed in discourse and testimonial settings. They may often *not be* consciously recognised by the scientists on

[136] For one view on this, see Philip Kitcher (2001). Kitcher argues that, under certain social and epistemic conditions, our quest should truth should not proceed in cases where it is foreseeable that inquiry would cause the lives of those who have historically lived the worst sorts of lives to go worse than they otherwise would have.
[137] Kathleen Okruhlik (1994).
[138] Thomas Kuhn (1977). It has been questioned as to why all scientists must share these particular standards, which is how Kuhn presents them. I will take it however that these values, if not universal, are at least widespread.

whom they operate. This leads to natural problems for the identification of potential sources of bias by the laity, and indeed by other experts. While not severely undermining science's place as the preeminent form of inquiry available to us, they help to form a more developed and measured understanding of it, and so help to act against misplaced confidence and/or overconfidence in scientific domains, in turn providing a platform for epistemic egalitarianism in that respect. We can easily see how class, gender, ethical, and ideological interests can come to influence scientific methodology and discourse. While this need not be greatly damaging to the veracity of science, as Galton's example shows, it does hold that possibility inherent to it, and hence must be guarded against.

We may choose to ask, as feminist philosophers have, whether the fact that science is predominantly done by men carries any significance for the methods and findings of science. Can we expect men to be as concerned about women's health as women are? Does gender influence scientific problem choice? In what ways do economic and political interests shape or influence scientific funding and research? For instance, can the research of scientists who work in 'Big Tobacco' be trusted? In what ways do pragmatic considerations shape scientific inquiry? How does theory and ideology influence scientific research and methods? For example, what part if any did a mechanistic or religious world view play in justifying the live dissections of dogs and other animals by early enlightenment-era scientists?[139] These, and many other questions of this type, are important for us to ask, or at least to be prepared to ask.

To summarise these points, we have seen that an important part of the motivation people have in engaging in pursuits, including truth-orientated pursuits, is self-interested, and that while those self-serving motivations can in some ways be notably beneficial to the epistemic community, it can in others ways be undermining of truth inquiry at the level of both a) the individual, and b) the group. We have seen that c) it is possible for the shared theoretical commitments of groups of experts, which are closely tied to the source of their understanding and their claims to epistemic authority, to be inhibitive or obstructive of a reliable veritistic connection. We have seen that d) forms of expertise, such as in the sciences, are modes of truth inquiry very much embedded within a particular historical and social context, and that this context can act as a notable epistemic influence upon scientific activity. And lastly we have seen that e) discourse amongst experts can in cases clearly reveal the tacit operating ideological and political commitments that are seldom explicitly advised to recipients' of testimony in advance. One final area of concern treated here is how linguistic practices might shape epistemic relationships.

3.4 Problems of Intellectual Authority: A Linguistic Analysis

> I therefore read the papers with some hope that they would help me "transcend" these limitations [of rationality, logic, science], or perhaps suggest an entirely different course. I'm afraid I was disappointed. Admittedly, that may be my own limitation. Quite regularly, "my eyes glaze over" when I read polysyllabic discourse on the themes of poststructuralism and postmodernism; what I understand is largely truism or error, but that is only a fraction of the total word count.

[139] For an overview and discussion of the influences on vivisection and animal experimentation in England, see Anita Guerrini (1989); for discussion of the role that Cartesian philosophy played in justifying this kind of scientific experimentation, see Peter Harrison (1992).

--Noam Chomksy, 'Rationality/Science'.

Philosophers have long found an interest in how language influences what we see, think, and do.[140] Feminist philosophers, for example, have been particularly vocal in urging us to pay greater attention to the way in which certain uses of words may unconsciously reinforce gendered stereotypes and shape social power structures, in doing so drawing a strong social and moral dimension to something as naïvely benign as 'mere words'. More generally though the way in which language is used by speakers is of central significance to the issue of epistemic relationships and epistemic dependence, and so is worthy of our attention. Indeed, so great is the power of language to influence epistemic relationships that I view this kind of discussion as being of the utmost importance to any encompassing veritistic social epistemology.

As all within it will be aware, one broad feature of academic discourse is that it is often engaged in using highly technical and specialised language that does not allow for an easy external assessment by those unaccustomed to it. Such discourse tends to be elevated from normal or traditional forms of speech, incorporating many unfamiliar terms, and assuming amongst the audience a much more extensive linguistic and subject background knowledge than could be expected otherwise. The justification one may most often find for these sorts of linguistic practices is that where the clear identification of certain concepts or sets of concepts is of some central importance, as it is in academia, that language use in this way is justified to the extent in which it aids precision. Thus, we may use a more technical form of speech or manner of writing in order to advance conceptual accuracy. In other respects it might be said that this kind of linguistic unorthodoxy is forced by the very nature of the topic itself. Our commonplace language, it might be said, is at times inadequate to the task of representing the concepts and ideas that underlie some area of study (such as one may find claimed in areas of mysticism, theology, or metaphysics). And so, because the concepts there go beyond the natural or the usual, so too must our language. In either case, by specifically using language in order to most accurately pick out particular concepts as intended, conceptual error and ambiguity can be reduced or avoided. Where the conceptual accuracy is of a more mundane variety, however, and not forced by the subject, we can expect that issues of targeting apply. Academics within a particular sub-domain of expertise who direct their address at other experts within that specialised field may use language of a more technical nature and assume a greater level of background knowledge than if their paper was also or only intended for experts of a more general domain. The same applies again if it is intended to be available to those in even broader areas of expertise, but still within the circle of academia. So, for example, if a philosopher who specialises in the philosophy of mind writes a technical paper on some nuanced point within that topic intended only for others within that particular sub-domain, we can expect that it is legitimate for the paper to be of a more complicated linguistic form than if he or she had intended for it to be read by philosophers more widely. And if he or she intends for the paper to be available to academics from diverse backgrounds of study, such as from psychology or mathematics, the kinds of things the paper can legitimately assume to be within the background knowledge of readers has to be calibrated to suit the audience

[140] And as was seen in chapter one, how it can even influence what we remember.

again. If the intended audience additionally includes the public, then the same principle of calibration applies again.

We may recognise further that non-standard word use can arise out of the aesthetic intentions of the author. Writers can seek to use language in non-standard ways in order to advance the literary quality of a work, and against this there can be little objection, except to express concerns that in doing so the availability of the content and conceptual accuracy is not undesirably undermined. Other reasons might see language as playing a demarcation function, differentiating one area of expertise and group of experts from another, and thus aiding identification.

Aside from any other reasons that can be provided, there remains one notable reason for engaging in unusual forms of speech and writing, and it is one motivated by psychological factors. Some deployments of language perform in fundamentally exclusionary ways, as a bulwark to out-group accessibility, and so act to insulate a work from external criticism. Language can serve to laboriously tie up an audience's mental energy in the antecedent task of identifying what concepts and meanings underlie the semantic knot created, and through this diversion of mental resources give obstructive pause to the assessment of the reasonable merits of the claims made.[141] This interferes with those seeking to analyse an argument on the basis of its content and argumentative structure before accepting its conclusions, and therein encourages deference to authority insofar as anything is understood at all. This is all the more significant in light of the fact that psychological research does suggest that when forming beliefs we weigh up accuracy goals against the costs in time and energy of being careful and diligent in our considerations.[142] The result is that where the costs of examining some work reach too high, because of the difficulties imposed in just understanding what is being expressed, we may relinquish or lower our demands in the face of putative authority, and so be helped towards believing deferentially.

In a mutual vein, speaking or writing in unnecessarily arcane language promotes uncertainty proportional to the unfamiliarity of the language to the hearer and the difficulty of understanding it, which in turn impedes our critical capacities through lowering our self-confidence to be critical. This unceremoniously helps us towards believing on the basis of our humility that the fault for the failure in comprehension must rest within ourselves, rather than the speaker.[143] To promote the inference that genuinely meaningful content exists beyond the menagerie of words, a writer may in effect "[bully] the reader into granting that, since one cannot figure out what is going on, there must be something significant going on, some complexity of thought, where in reality there are often familiar or even shopworn notions, addressed too simply and too casually to add any new dimension of understanding."[144] Where the language is sufficiently complex this can be true even for those with some significant background in the discipline in question.

[141] Martha Nussbaum (2000).
[142] See Ziva Kunda (1990).
[143] Denis Dutton (1999).
[144] Martha Nussbaum (2000).

Lastly, but simultaneously, excessive representational complexity can promote the appearance of conceptual complexity, and therein promote the appearances of intellectual merit. If the ideas involved are essentially simple or basic, and easily graspable by most, then there is no particular merit in having had them. But by wilfully and creatively imposing linguistic difficulty onto ideas by dressing them in grandiose and mystifying language, a complexity can be given to them that is not of their own. So entangled, this artificial complexity can serve to help falsely venerate the intellectual. It can be a way for those inclined to set up the appropriate conditions in order to give a false appearance to their own intellectual capacities and produce, while simultaneously drawing attention to both. This too encourages deferential acts.

Of significance is not just how what is said is in fact said, but additionally what remains unsaid and is left out. Withholding, failing to provide, or too easily assuming as known information which an argument or discussion depends upon further accentuates the knowledge gap between speaker and hearer, and in this can aid in promoting deference. If a thinker assumes as properly known some point *x* amongst the audience, and yet *x* is not known by most, then an audience's reflections upon that knowledge space may inspire a sense of anxiety within them in relation to the perceived epistemic inequity. Further, by not saying enough to allow an audience (or parts of an audience) full independent access to the substance of the content, deference is promoted by limiting the ability of others to believe independently on the merits of the argument alone, without any testimonial reliance. The costs of ensuring that adequate information is provided however must be weighed against the pragmatic limitations of spelling out in full detail every item which might be otherwise expected to be known by the majority of a target audience.

Given these possibilities, the following picture emerges: intellectual insecurity and dishonesty can lead to the selection of certain modes of language use intended to a) aid in the concealment of absurd, vapid, or otherwise exceptionable content, b) insulate the work from external sources of criticism, and c) tie up the recipients mental resources in trivial antecedent linguistic tasks, thus delaying and making more costly critical assessment. While obstructing our ability to be critical, it may further d) lower our willingness to be critical by setting up the conditions in which we may judge ourselves at fault for an absence of understanding. In the absence of clear understanding it can e) promote the inference that genuine content exists, create and enhance a perception of an asymmetry of knowledge between speaker and hearer, and aid in the veneration of the thinker by encouraging linguistic complexity to be confused with conceptual complexity, all of which in turn promotes epistemic trust and deferential acts. These factors taken together constitute a significant negative force of influence upon epistemic relationships. Using language in this way is fundamentally exploitative of the epistemic and social conditions under which testimony occurs, and acts to undermine the integrity and honesty of the situation in order to advance academic, social, and/or psychological ends.

This is, alas, nothing new. Language in its negative employment has been recognised by thinkers for some length of time, such as in the writings of Voltaire, Goethe, and most

especially in Schopenhauer.[145] Schopenhauer wrote with bitter and seething disdain for the way in which some thinkers subverted language for nefarious ends. He chastised the 'colossal mystification' of some of Europe's most admired philosophers, saving his most savage derision for him whom he saw as the greatest abuser of language in this way, his contemporary and adversary Hegel. Incensed, he wrote

> …they try to make the reader believe that their thoughts have gone much further and deeper than is really the case. They say what they have to say in long sentences that wind about in a forced and unnatural way; they coin new words and write prolix periods which go round and round the thought and wrap it up in a sort of disguise. They tremble between the two separate aims of communicating what they want to say and of concealing it.[146]

He condemned

> that mode of concealing the most awful poverty of thought under a babble of inexhaustible chatter that resembles a clacking mill and is just as stupefying: one may read for hours together without getting hold of a single clearly defined and definite idea.[147]

In contemporary times postmodernism has drawn great ire from critics for its resiliently incomprehensible manner, its ambiguity, and its apparent celebration of commonplace or mundane truths. Noam Chomsky has criticised postmodernism and poststructuralism along these lines by collecting examples of what he takes to be reflections of the ordinary nature of so much of what is expressed (so far as it is comprehensible to him), the falsehood or inanity of others, and the sheer impenetrability of the language used there generally.[148] While Chomsky does not explore the particular psychological and epistemological benefits available to those engaged in this way of communicating, Martha Nussbaum has offered considerations to that effect. In an article taking aim at Judith Butler as something of a model of language abuse,[149] she argues (alongside some points mentioned earlier) that obscurantist language can act as something of an intellectual watermark on ideas, and in this can be something of an academic equivalent of scent marking. "When ideas are stated clearly… they may be detached from their author: one can take them away and pursue them on one's own. When they remain mysterious (indeed, when they are not quite asserted), one remains dependent on the originating authority."[150] We may then place questions before writers who seem to us to invent superfluous neologisms, and multiply the complexities and depth of language unnecessarily, for the speculative purpose of influencing the form of future discourse for their benefit.

---

[145] Goethe has been attributed as saying that "when ideas fail, words come in very handy", while Voltaire wrote that "a profusion of words is the great vice of all our modern philosophers and anti-philosophers" (1764, section titled 'style'). Voltaire also quotes Horace's saying that "simplicity relieves grandeur".
[146] Arthur Schopenhauer (2004a).
[147] Arthur Schopenhauer (2004b).
[148] Noam Chomsky (1995).
[149] Martha Nussbaum (2000). Richard Dawkins does similarly in his (1998).
[150] Ibid.

We should recognise also that obfuscation is not purely one way in terms of the benefits it can deliver. While on one hand authorities are able to reap rewards for communicating in difficult ways, it is also true that *understanding* a difficult work provides readers with certain benefits, for in this they can show themselves to be among those who are intellectually sophisticated enough to understand such a difficult and putatively deeply sighted work. If the form of a work can act to limit the number of those who can have intelligible access to it, and the content within is held out to be of significance, then to understand it is to be the bearer of special knowledge and intellectual sophistication that others may not easily come into on their own. We can see well why some might then protect and defend certain putative authorities from external criticism along epistemic and linguistic lines, because it is precisely on the soundness of those authorities (conceptually, linguistically, veridically) that the value of their associated achievements in understanding them hangs.

That effect may be reinforced by a need within us to justify to ourselves how we have spent our time and energy. Pointlessly expending great effort on cryptically hidden ordinary truths and falsehoods is not, other things being equal, a desirable outcome for any of us to arrive at. After having striven for some length at insight we may be subconsciously moved to think more positively of a work, simply in order for that effort to not have been wasted on nonsense.

I find it doubtful that encouraging more favourable judgments from readers explicitly forms part of the consciously motivating reasons for most cases of obfuscation, for it seems unlikely that any putative expert's thoughts could deliberatively extend that far ahead before choosing a style of writing or speaking. It seems then more likely that it proceeds from a more subliminal, unconscious choice. While some may write *primarily* in order to take advantage of language in these sorts of ways, for others these kinds of aims form background elements in the manner of their communicating. That is, these sorts of considerations are (weak or strong) influences on the linguistic character of a work, rather than the driving motivation for it. One further conciliatory note I would make is that due to an effect of the language, those who engage in language abuse may be none the wiser about it. It is quite possible that thinkers may partly mistake their 'difficult to understand' manner as an indication of the profundity of their thinking, or even its truth, and in this way are deceived about it just as others are. Self-deception can be a powerful thing, and the appearance of genius in one's own writing may retard self-criticism just as readily as it can act as a block to external criticism.

In any case, it is arguable that more arcane language exists in various areas of academia than can be legitimately defended on the basis of conceptual or aesthetic goals. What these reflections provide us is an opportunity for re-emphasis upon the intellectual honesty and courage to expose one's ideas for what they are, and to refrain from dressing them up in 'anxiety inspiring', linguistic labyrinths. What is called for is an on-going reaffirmation of an academic attitude and culture that repudiates works which lean too far towards obfuscation, technicality, vagueness, and ambiguity, for no perceivable legitimate purpose or gain.

One final point to place this discussion in perspective: we can find the utility and practice of unconventional language use in areas far wider than that of academic discourse. One of the quite notable things that the Church of Scientology founder L. Ron Hubbard took measure to do when introducing his teachings was to come up with a great many new words, abbreviations, and acronyms dedicated to representing his ideas. With Scientology, as with many kinds of philosophy, one is not simply taught a new way of thinking but also a quite new *way of speaking*. Many of the new terms introduced however appear to be little other than redefinitions of pre-existing words or replacements of existing words which already adequately function to convey the intended meaning.[151] The effects of this are potentially numerous, and accord with what has been outlined so far: the promotion of the appearance of an epistemic authority, the tying of claims to certain restricted epistemic sources, the obstruction of critical reasoning through the diversion of cognitive energy (including retraining and modifying one's linguistic knowledge), a misleading imposition upon the conceptual difficulty of the system of ideas through semantic manipulation and alteration, and so on.

3.5 A Questioning of the Questioning

I consider here three anticipated objections in relation to the suggestion that there is a place for a social critical examination of the historical, psychological, and epistemic characters of epistemic relationships and figures. The first suggests that such a project would not be sufficiently worthwhile to justify the harms it would cause. Philip Kitcher, for one, has argued that truth may not be best considered as an absolute good, and that *some* truths should potentially not be pursued or made known in the case that they would have certain clearly adverse effects for society.[152] It could be argued in a similar vein that the number of undesirable consequences that would develop from a critique of epistemic authorities would make such a project undesirable. This argument we might term the 'argument from insufficient utility'. In what possible ways might harms come about? One way could be that it would cause a sceptical overreaction, causing more doubt in epistemic authority than could be a) justified by the facts, or b) desirable in light of the efficient dissemination of true beliefs. Another way might be that society could lose accessibility to many previously valuable social 'role models' as a result of famous figures being seen in less idealized ways.

In response to the first point, I can think of no compelling reason to grant significant weight to that possibility. We are too deeply invested and indebted to epistemic authority, and ever increasingly, for a widespread radical scepticism towards ESP's to be either

[151] See for example the Church of Scientology's Online Glossary of Terms.

[152] Philip Kitcher (2001). He illustrates the non-absolute utility of truth by using the hypothetical example of the discovery that cooking certain ingredients in exactly the right proportions could create an explosive material far in excess of the power of the most powerful nuclear weaponry. Virtually all would agree that this not be a truth to be released into the public body of knowledge, given the recipe for disaster such a truth would cater to (no pun intended). In real world examples, truths about differences between men and women, or different racial groups, could be argued to be not advisably made widely known in the case where their being known would bring about grievous social harms.

appealing or a rational development. The mere fact that many types of experts are fallible and subject to prejudices does not (alone) warrant an extreme sceptical response. If a minority draw these conclusions then this is rather unfortunate, but alone does not inspire substantial worry. I don't see a widespread scepticism as a realistic development for the majority, especially when it is made clear as to the benefits of epistemic interdependence, and so do not see this as being something worthwhile to anchor discussion on. Certainly, nothing about any of the preceding discussion impels in me a sense that my many epistemic dependencies must be cast into a universal situation of doubt, and I suppose myself not to be exceptional in that.

The second claim is that if some thinkers are upheld as the paragons of scientific or intellectual pursuit, then any bringing to light of their errors undermines that image of them insofar as they provide an ideal that others may aim towards. This is a slightly curious objection, for it seems to suggest that we cannot help but think only in 'black and white' dimensions. The importance of role models can be greatly significant to many, as they can help us break free of the arbitrary or artificial limitations that are placed upon us (by ourselves or by others), and aid us to move towards higher and better things by setting powerful, resounding examples. Yet I take it to be a simple matter that we do not think radically less of Newton's truly remarkable scientific achievements because he was not a model of personal virtue. Nor does the fact that he was a devout Christian seem to unravel the widely atheistic or agnostic academic community in their willingness to appreciate the successes of his intellect and scientific work. Nor does the fact that he devoted massive tracts of time to distinctly pseudo-scientific subjects such as alchemy and biblical prophesy render meaningless his scientific successes. This is just one example of course, but the same holds more generally. I see no particular reason to think that a heightened understanding of exactly where great thinkers have historically gone wrong - according to our modern eyes - is entirely devastating to the possibility of their being valued or appreciated by us. We can be rightly moved by what virtues they honestly bear and exhibit in their achievements, and take *aspects* of them if we so choose it as models for us to set ourselves towards. That indeed is the wisest position to take in any case.

A second avenue of criticism suggests that this project is disingenuously conceived, and derives its motivation not from the hope of a realisation of higher social and epistemic goods, but from the gratuitous satisfaction of base desires. Let us call this the 'argument from schadenfruedian underpinnings'. It argues that the scrutiny of our most cherished and respected thinkers proceeds much more easily and probably from lower intentions than virtuous ones, and in this comes about in a compromised way. This concern is legitimate in the sense that it is indeed possible for people to derive a certain pleasure in the public depreciation of others. However no suggestion is made here that epistemically significant figures should be held to unreasonable standards of conduct or belief in order that they provide examples useful in this project, or that their venerable achievements are to be undermined and discredited in view of their humanness. There can be little else said on this matter except to deny it. We may certainly have an anthropological pleasure and intrigue in the various items of belief people have attended to over the course of history,

but I don't see this as undermining the merits of what is proposed – for there clearly are epistemic goods possible.

Connected to this, a third objection is more cautionary, and insists that the faults of epistemic authorities must only ever be considered when located within the social and political contexts of their mode of thinking. It argues that by not giving explanations in terms of social, historical, and political influences, that ESP's are bound to be judged unfairly.

Let us take as our working example the case of David Hume's judgments against black skinned people. It is surely worth knowing, in terms of seeing how or why Hume concluded what he did, that the breadth of his personal experiences with people of darker skin was probably limited, and hence that he had fewer opportunities in terms of direct observations to be provided with evidence that disconfirmed his seemingly racially prejudicial judgments. Also of relevance is that such views were not uncommon within the social fabric of his day, and so would likely have been seeded and supported by the social diffusion of ideas. Of course it cannot be said that greater personal exposure to black people *would* have caused him to have abandoned his prejudices, since that would make a number of assumptions, including that he would not have found ways to dismiss conflicting evidence.

In light of this understanding we are immediately called by Hume's example to examine the ways in which we may be similarly influenced in our thinking. In what ways are our own beliefs dependent upon our (limited) experiences? In what ways is our thinking influenced by the character of our social environment and its background beliefs and values? In what ways are we also inclined towards resisting giving up beliefs in the face of recalcitrant evidence? These are greatly important questions for us to ask; equally important is to know simply *that* a great mind (as Hume surely possessed) is not enough in itself to protect against these epistemic concerns. I entirely agree that context should play an important role in an analysis of historical thinking, not so much because this can serve to balance judgments about the individuals in question themselves (although this is valuable), but because it offers the possibility of yielding so much more that is of substantial relevance to the understanding that we seek for ourselves.

Accordingly, nothing about what is proposed underplays the value of any of these broader considerations. An exploration of the philosophical, ethical, and factual confusions of historical thinkers has value in so far as they provide examples from which we may come to further see how in a number of cases the 'intellectual leaders of mankind' (to use Popper's term) failed to resist or see its way past intrinsic and extrinsic sources of error. By such an examination we are invited to see how integral our background beliefs and the conditions of our observations are to how we deal with conflicting evidence and reason about the world and others. All of this draws us into deeper contemplation of our own somewhat precarious epistemic situation within the social world. To claim with frustration that such enterprises are anachronistic is to miss the valuable point about it all, which is that it is *precisely in* considering these thinkers *both* within and outside of their particular historical and social contexts that insights are

made available to us. The mere fact of error alone has value just the same as a broader understanding of how the error occurred, and both aspects should be pursued where available. While certain costs are involved in this, it's anticipated on-balance utility outweighs those concerns.

3.6 An Ancient Project Revisited

With these considerations in mind, it is well to see that our confidence in experts should be measured against our understandings of the ways in which epistemic relationships are shaped along social, psychological, epistemic, historical, and theoretical dimensions. Even the accomplished and venerable pillars of intellectualism are subject to many of the same truth undermining features of human cognition as can be found in all places, in addition to being vulnerable to the features specific to the very condition of expertise itself. In the course of discussion we have seen that there are a good number of reasons to warn against an unreflective and uncritical dependence upon various sources of expertise. The social illumination of these realities would be a significant first step towards the reduction of the moral, veridical, and prudential problems associated with our universal and necessary dependence upon forms of expertise.

This task is little different from that which Socrates took for himself, some 2500 years ago. He determined himself to play a role in shaping the thinking of his fellow Athenians, acting as the 'gadfly' to their sleeping epistemic scrutiny and responsibility. For so many of the Socratean interlocutors this experience was a distinctly uncomfortable and even painful one, as changing ones beliefs and broadening ones awareness sometimes or often is. As a consideration this is not trivial by any means, and it does weigh upon the expectations that we can have for such a project. Nonetheless, with Socrates, and according to my best judgment, it stands that the value of encouraging reflection and understanding outweighs the value of avoiding what kinds of pains may be naturally inherent in doing so.

We are then to ask how we are to proceed with regards to epistemic authority. Some of the answer to this question will rest upon the epistemic nature of testimony itself.

On the Conditions of Testimony

4.1 Whom to Trust?

What we mean by the term 'expert' is a good place to start for a discussion of expertise and deference. Expertise may be defined in a comparative manner, where an expert is any person who comparatively possesses more knowledge in a particular given subject than nearly all others.[153] Given this definition, an expert is one who knows more than others do about some subject S. But how are those who know little of a subject to recognise those who know much of it? That is, how can we recognise other people's knowledge *as* knowledge when we ourselves do not possess it as such? This is a fundamental problem that bears strongly upon our ability to assign authoritative status to others at all.

Alvin Goldman has suggested that much of the answer is provided by taking note of the temporal dimension of it all. While we ourselves might not be able to see another's claim *x* as knowledge at T1, we might at T2 possess this same knowledge, and so retrospectively be able to see that they did possess knowledge (of *x*) at T1. That is to say that we cannot see what others know before we know it ourselves, but we can see what others *knew* before we knew it. This provides us some ground on which to assign a level of epistemic authority to others, and this is given additional strength in the case that there was nothing at T1 that we knew which the experts did not know in the subject in question. An asymmetry of knowledge between speaker and hearer, then, supports us in justifiably assigning expertise to people. From this Goldman suggests that we can inductively conclude that some person's body of knowledge in a given area exceeds our own while not actually knowing what it is that they know which we do not.[154]

How might this work out in practice? One way is through observation: we can observe new facts about the world, or contrary facts to what we previously believed, and see that others knew these facts before us. Another is argumentative: we can be shown that something is true (that we were previously either agnostic on or denying of) through argument on the basis of what we already believe to be true, and so see that others knew what we now accept in a given area before our knowing it. If this happens a number of times we are provided good reasons to think that the people in question have a level of knowledge that meets the previously given definition of expertise.

One possibility allowed for is our deferral to experts on the matter of who is an expert, a process which would offer to greatly increase the efficiency with which we are able assign authority to others. Suppose that *A* is an authority on who else is an authority in some area (e.g. scientific knowledge). If we are able to sufficiently determine that they have this knowledge, we are able to classify as authorities those whom they specify to us as such. While particular instances of these could be defeated by future experiences, which if numerous enough could then call into question our judging of *A* as authoritative

[153] Alvin Goldman (1999), p. 268. He suggests that they also have an appropriate intellectual skill set related to their expertise.
[154] Ibid, pp. 268-269.

in this way in the first place, but failing such circumstances we gain access to a whole plethora of sources of authority. But how can it be that *their* knowing can become *our* knowing? What can epistemically entitle us to simply accept the beliefs of others as our own? Is this in fact what we do?

4.2 A Distinction in the Justification of Testimony

Much of the focus in epistemology has traditionally been upon the thinker alone in the world, and has sought answers as to how it is that they can claim knowledge on an individual basis. More recently epistemologists have become much more appreciative of the fact that we believe many more things as a result of being presented them as true by others, than as a result of our having independently come by these beliefs ourselves. In view of this we find ourselves indebted to and dependent upon testimony for a great amount of what we can claim to know. But of the multitude of beliefs we obtain from others, we have the time, capacity, and resources to independently directly verify the truth of only a vanishingly small number of them. It would seem then that were it the case that each of us were an epistemic island, isolated from the abundance of knowledge made possible by our social existence, then the sum total of what any one person could come to know in their lifetimes would be vastly diminished. This invariably raises questions regarding the specific nature of the way in which we are invested in accepting and taking as true what others say. From where does our confidence arise in the truth of the tellings of others?

There are two broad ways in which to go about justifying testimonial practices. The first is the antireductionist route, which is a position historically identified with Thomas Reid. Reid thought that divine providence saw us made such that we are tempered to tell the truth and to believe the reports of others, and we may do so unless we are able to find sufficient reasons not to. There are several species of antireductionism, and although I will give time only to antireductionist foundationalism (which in places seems spoken of almost as though it is antireductionism entire) the reader should be aware that it alone does not exhaust all antireductionist possibilities.[155] As a position foundationalist antireductionism attends that we are simply entitled, as a brute epistemic fact about us, to believe the tellings of others in basic cases, except where we have sufficient reason not to do so. In much the same way as many hold that we are entitled as a basic epistemic right to think that beliefs acquired through perception or from memory reflect the way the world is (or was), so too is the testimony of others to be considered a similarly fundamental way of knowing.[156] In this traditional antireductionist account background beliefs play a largely negative role. That is, they may stop belief where we otherwise would "just believe".[157] All antireductionist principles are not equal however, and some have been formulated in a way that is responsive to uneasiness about brute entitlements.

[155] See Goldman (1999), pp. 128-130, for an overview of four antireductionist models of testimony.
[156] For an antireductionist conception of testimony as "conditionally basic", see Robert Audi (2004), pp. 18-34.
[157] Paul Faulkner (2002), p. 22, citing Audi.

Peter Graham, for example, offers a sensitive 'liberal' (i.e. antireductionist) principle in the following way:

> If a subject *S* comprehends both the force and content of a telling by another that *p*, and if this causes or sustains in the normal way *S*'s belief that *p*, then this confers *prima facie pro tanto* justification on *S*'s belief that *p*.[158]

We can see in this version that the justification provided is somewhat tentative and qualified – the justification given is *prima facie* and *pro tanto*, where *prima facie* justification means that the justification is defeasible, and *pro tanto* that the justification may fall short of being sufficient justification. In any case, in assigning fundamental rather than derivative trust to people, antireductionist accounts are thought to offer certain advantages over reductionist accounts. Perhaps most significant among these is that they would seem to offer a way around what is held by some to be the impossible task of actually finding independent reasons for the acceptance of so many of cases of testimony.

Standing opposite to this approach is the reductionist account, a position historically associated with David Hume, who is taken to be rather unique for thinking about testimony in this way.[159] Hume, like Reid, noted the necessity and propensity for the acceptance of testimony in the course of life. "There is no species of reasoning more common, more useful, and even necessary to human life", said Hume, "than that which is derived from the testimony of men".[160] But while holding the testimony of others as essential to us, Hume tempered considerations by casting scrutiny onto our reliance upon it. "[N]o weakness of human nature is more universal and conspicuous than what we commonly call CREDULITY, or a too easy faith in the testimony of others…"[161] Faulker writes that for Hume "belief is more properly an act of the sensitive, than of the cognitive part of our natures", and that more often than not the practice of testimony "commands our assent beyond what experience will justify..."[162]

Hume's reductionist position was that "we are justified in believing testimony only because we have observed past conjunctions between reports and reported facts", and "that the conjunctions we observe are between types of testimony and the truth of instances of these types".[163] Hume thus saw the justification of testimony as occurring through the observation of how those types of testimony relate to the state of affairs in the world; by the observance of the veridical conjunction between a testimony type and fact. Hence, the reductionist route does not allow for testimony to be taken as a properly basic way of knowing. In contrast to the antireductionist account, the merely intelligible expression of a proposition by a speaker provides no justification or entitlement *in itself* for the hearer to accept it as true. Our acceptance of testimony, for the reductionist, is tied to more fundamental ways of knowing.

---

[158] Peter J. Graham (2006).
[159] Paul Faulkner (1998).
[160] Ibid, citing Coady, who cites David Hume, 'An Enquiry Concerning Human Understanding'.
[161] Ibid.
[162] Paul Faulkner (1998).
[163] Ibid.

4.3 Problems on Either Side

An investigation in the epistemology of testimony reveals numerous complexities and complications, and a number of criticisms have been launched against both conceptions, of which I can only survey some important ones. Against reductionism it has been alleged that the observation of the concordance between testimony and truth is ineliminably bound up in testimonial practices. On this account, testimony cannot be extricated from the very things supposedly used in order to justify testimonial practices,[164] for in order to gauge the truth of so many instances of testimony we will invariably end up relying upon testimony in some way (at some level, in some manner) in order to do it. The circularity of that process is taken to invalidate it. How so? We can easily see this by looking at the example of sense perception: what justifies us in believing that the beliefs we acquire by visual means reliably reflect the states of affairs of the world? If for example we were to see an apple on a table in front of us, what would justify us in thinking that it actually exists? One way might be to argue in this fashion:

> On previous occasion O_1, person S_1 perceptually formed the belief that P_1, and P_1 was true
> On previous occasion O_2, person S_2 perceptually formed the belief that P_2, and P_2 was true
> …
> Therefore, sense perception is a reliable source of belief.

Such an appeal would be a "track-record" approach to justifying visual perception.[165] However such a method of justification tacitly appeals to visual perception in justifying the reliability of visual perception, and hence must therefore be circular in justification. For any person sceptical of the reliability of visually acquired beliefs such an argument would provide little force against them, as it seeks to justify the thing under scrutiny by appealing to that very thing. An alternative response would be to suggest that we could at least go over to the table and pick up the apple, and so gain tactile information about the apple that agrees with the visual information. This second attempt then would seek to justify visually acquired beliefs with the veracity of our tactile senses. This is a perfectly acceptable way of giving justification for our visual faculties, but we can of course inquire still further into what reasons there are that could justify our belief that beliefs acquired through touch accurately reflect the world. Making reference here to visual perception will not do, as we have enlisted tactile sense to justify perception in the first place.

At the widest possible juncture of epistemic justification lies the 'superpractice' of human cognition, and as it is clear to see here that any attempt to justify human cognition will necessarily have to utilise human cognition in doing so, thus rendering any such attempt argumentatively futile. It is in fact necessary that such circularity enters, for in any case

[164] Alvin Goldman (1999), p. 127.
[165] Ibid, pp. 83, citing Alston (many).

where the number of practices is finite it will be impossible to endlessly justify one practice by making reference to another not already referenced, or without invariably making reference back to itself. Similarly so for testimony: in observing the way in which testimony relates to truth we must surely rely on testimony along the way in making those observations (for instance, our perceptual judgments may involve classifications acquired testimonially, about what this or that thing being seen *is*), which is to tacitly bring it into the argumentation that seeks to justify it. One can find then a certain appeal in antireductionist models, which side-step this problem of circularity altogether by seeking it outside of other foundational groundings.

However, if it is true that testimony is tacitly appealed to in justifying reductionism, we need not necessarily think of this as being absolutely devastating to reductionist models. Goldman has argued that the coherence of two visual reports about an object is at least *some* evidence that the beliefs acquired about it on the basis of perception are reliable, as it is not obviously necessary for any two perception-based reports to have to be in agreement with one another. In his words, the one is not destined to "rubber stamp" the other. We could, for instance, look at an object on a table and see it as an apple, but on getting closer see in fact that it was something else, and that we were mistaken.[166] If Goldman is correct then we could apply the same point to the case of testimony. That is, we would be entitled to say that because it is in no way necessary for a fully coherent picture of things to arise from testimony, that a coherent picture does tend to indicate *something* on the side of the reliability of testimony in a way that (to some extent) dissipates circularity concerns. If that is the case, then the problem of the circularity of reductionism may be seen as less significant than it at first might appear. It would no longer be a purely reductionist picture of things, however.

While testimonial reports may not need to be universally in accordance with each other, it has been argued that they must coincide at least a lot of the time. This line of attack upon reductionism was explored by C. A. J. Coady, who argued that since in the case of reductionism the reliability of testimony is supposedly checked against reality, in theory it should have been possible to find that there was no connection between testimony and truth at all. However Coady argues that this kind of discovery is *a priori* impossible, since in that case there would be nothing like testimonial practices to observe. The inference then is that reductionism must be mistaken, and that testimony is broadly necessarily reliable (and so our confidence in testimony justified *a priori*).

It is an interesting argument against reductionism, but a number of worries present themselves in reflection upon it. In a general line of criticism Graham argues that Coady's argument fails to convince most essentially because of the intuitive implausibility that what people say is *necessarily* more likely to be true than false. It might be that *as a matter of fact* that we tell the truth more often than not, but the claimed necessity of our having done so seems to be a deeply troubled proposition.[167] Graham is correct, but to actually see why I will develop an analogy with the anthropic principle.

---

[166] Ibid, p. 85.
[167] Peter J Graham (2000). See this for several other specific criticisms of Coady's argument.

Coady's argument, as presented, is essentially a transcendental argument for the reliability of testimonial practices. A transcendental argument is one that attempts to argue from the uncontested nature of certain propositions to the *a priori* necessity of others. The anthropic principle is an example of a transcendental argument. In one manifestation the anthropic principle argues that since if the universe was not sufficiently complex to support life there could not have been life within it to observe this fact about it, we must say that *we could only ever* have observed the universe as being sufficiently complex to support life.[168] As such, we really should not be surprised to find the universe life-supporting. In the hands of some commentators this sort of reasoning can be presented as though it does explanatory work for the apparently unlikely fact that the universe is the way that it is, 'tuned' towards the existence of life.[169]

The extended reasoning taken towards the explanatory value of the anthropic principle is highly questionable, and the error of it can be brought to the surface using a well known counter-argument involving riflemen. Suppose that we are lined up in front of a firing squad set to be executed. The order is given by the sergeant, the crack of the guns is heard, and then... nothing. Finding ourselves alive, we react with complete surprise – to our delight and temporary relief all of the marksman's bullets have apparently miraculously missed their intended target. The question then immediately before us given this strange and unexpected turn of events is 'how did this happen?' We can all agree that the appropriate answer to this question would not be 'because if the bullets did not miss, we would not be able to ask why it was that we are alive'. Although that is most certainly quite true, and *necessarily* so, it is a truth that applies to the wrong level of description than the one which we want. We want to know what went wrong with the sights on the guns or the vision of the executioners, the bullets, or some other explanation of this type.[170] It is this same case in anthropic cosmology. Although it is true that we could only ever have observed a 'goldilocks' universe - one that is just right for intelligent life in a wide number of fundamental ways - this places no necessity on the universe to be as such. The appropriate question to ask is not 'why do we see the world as being the way that it is', but rather 'why is the world the way that it is, such as we see it'.

To return to Coady's argument, the apparent *necessity* of the reliability of testimony is similarly mistaken. Let us place the steps of each transcendental argument along side one another, for analogical comparison:

An observation is made:

1a) The universe is just right for life
1b) We have reliable testimonial practices

Transcendental reasoning is undertaken:

[168] The principle assumes that consciousness depends upon complexity.
[169] See Richard Dawkins (2006), pp. 134-151, for an account that is highly suggestive and even endorsing of the anthropic principle as doing some kind of deeper explanatory work.
[170] As Burton Richter (2006) wrote disparagingly of appeals to the anthropic principle, "the anthropic principle is an observation, not an explanation".

2a) If the universe was not just right for life, we would not be able to see the universe as not being just right for life.
2b) If testimonial practices were not reliable, we would not be able to find testimonial practices as unreliable.

And it is concluded therefore that:

3a) The universe is necessarily just right for life.
3b) Testimonial practices are necessarily reliable.

Although we can say that it is quite true that if people systematically told falsehoods at levels greater than truth that there would not be 'anything like reporting' going on in the linguistic community, this fact *when offered as an explanation* of why there are truth-telling practices and tendencies is one given at the wrong level of description. It simply does not follow from reporting practices requiring generally truthful participation, and from the existence of our testimonial practices, that testimony is necessarily generally reliable. As an account it confuses observation with explanation in the same way that offering the anthropic principle as an explanation for the specific features of the universe does. As such, Coady's argument in favour of antireductionism cannot serve as a sound justificatory basis for testimonial practices.[171]

Against reductionism, critics have argued that if it were true then we would seem to be confounded by what is taken to be an impossible epistemic burden of actually independently finding reasons to justify our trust of epistemic authorities. But is this true? Paul Faulkner has argued that the antireductionist assumption about the psychological reality of testimony is quite mistaken.[172] Irrespective of what *actually* justifies us in believing testimony, Faulkner argues that we at the very least think of ourselves as possessing a wide number of reasons for the acceptance of testimony.[173] That is, even if antireductionism is true and we are simply epistemically entitled to 'just believe', this does not match up with what occurs within us psychologically in relation to our confidence in cases of testimony. Faulkner lists among these reasons: an analysis of the way in which a particular instance of testimony is related to a type of testimony, and how that then relates to our background knowledge of the reliability of that testimony type, our beliefs as to the truth of the proposition and the costs of being wrong for the

[171] It has been argued that predation exerts a strong selection pressure for evolved mechanisms for the transmittance of warning alarms through signals, under the appropriate conditions. This seems much more relevant in relation to an explanation for shared truth-telling dispositions and practices. See Alvin Goldman (1999), pp. 106, citing Marc Hauser, 'The Evolution of Communication' (1996). See Graham's article for an exposition of more subtle ways in which Coady's argument fails to be convincing upon inspection.
[172] Paul Faulkner (2002).
[173] An antireductionist might respond by arguing that the antireductionist position makes no claims about what *we believe* justifies the practice of accepting testimony, rather than what in fact *does* justify testimonial acceptance. Continuing from this, it might be said that psychological considerations, while perhaps interesting, are also irrelevant to the issue. This is something Faulkner clearly acknowledges but sees as being of no relevance; the mistake of the antireductionist position is still in thinking that we are absent (psychologically) of reasons for accepting testimony.

speaker,[174] our knowledge of the particular situation in which testimony is being given, and (less validly, for Faulkner, because of false attribution bias) an analysis of the speakers character. Thus, he argues, it is not simply that our background beliefs play a "filtering role", stopping us from accepting testimony we are otherwise disposed to accept, for they in fact give us plenty of (at a minimum, psychologically) positive reasons to accept testimony from others.

This appears to be the case. Were I to telephone a representative of a business in order to find out how much money is owing on my account with them, while no *specific* thoughts as to why I should believe them may enter into my conscious thinking at that moment, I assuredly do otherwise possess them. The fact that thoughts towards justification may not be present at the precise relevant moment does not undermine their being held at all. This is no truer than it would be to suggest that because we seldom consciously think about what to do (or what we are doing) when driving a car, that we don't have any knowledge regarding how to drive a car at all. Elizabeth Fricker has supported this general point, arguing that even if antireductionism has truth to it, that in practice it is superseded by reductionist justification in the case of adults, as any 'normally knowledgeable adult' will possess manifold 'relevant background information' relating to the veridical status of sources of testimony.[175] On her view an antireductionist account of testimony is restricted in application to generally children and young adolescents, who as a result of their age are not in a position to possess an equivalent wealth of background beliefs that make justifying testimony reductively possible.

An argument offered by Graham in reply is that even if this is true, since much of the background knowledge that is brought to bear inferentially upon newly acquired beliefs will *itself* have been acquired in a non-reductive fashion (such as during adolescence), that antireductionism still plays an *indirect* part in the acceptance of new beliefs from epistemic authorities.[176] A point not made plainly clear however by Graham, but necessary for that argument to succeed, is that at least some of these beliefs acquired non-reductively are sustained in this very same fashion, and not later given reductionist justification independent of antireductionist justification. Nonetheless, whether it is possible for all beliefs acquired non-reductively to be given reductionist justification at some later time is questionable, and thus the possibility that antireductionist means of belief acquisition plays some kind of an indirect role in our beliefs must be taken seriously. Graham thus entertains and outlines the possibility of an 'over-determination'

[174] If, for instance, the costs of false-reporting are high for the speaker (as may often be the case for scientists), they will be much more likely to want to avoid being wrong, and this is taken psychologically by hearers as added reasons to be careful and to tell the truth. The cost of accepting a false belief for the *hearer* only has implications for the level of scrutiny needing to be met by them, offering nothing in relation to the probable truth-value of instances of testimony. That is, they may be more cautious about accepting important beliefs which if wrong would have significant negative results, and thus be more demanding in testimonial settings, but this adds nothing on behalf of the actual probable truth of whatever is being told.

[175] Peter J. Graham (2006), citing Elizabeth Fricker, 'Trusting Others in the Sciences: A Priori or Empirical Warrant?', *Studies in the History and Philosophy of Science,* 33 (2002).

[176] Peter J. Graham (2006), p. 90.

of justification for beliefs, such that both reductionist and antireductionist justifications could mutually co-exist for beliefs acquired in testimony.[177]

As mentioned, in reductionism the justification is said to exist between observances of instances of testimony and testimony type and the truth of the propositions involved. R. A. Naulty and P. J. Sheehan have argued however that in cases the classification of some report as belonging to a particular testimony type could be difficult.[178] They argue that testimonial categories could be expanded in some instances potentially indefinitely. For example, is the report of a delay to our aeroplane's departure time at the airport an example of an aeroplane type report? An airport type? A loudspeaker type? A delay type? Or some other kind? If each type required independent observation to establish a veridical connection, and indeed to a high standard of confidence, then the cognitive workload would seem to be unreachable. While this may be a possibility, also possible is that psychologically we largely intuitively or unconsciously class instances of testimony to testimony types. A lot of our thinking comes to us intuitively, and it seems *prima facie* plausible that in the vital social matter of assigning and organising trust in others that the categories of speech acts is assigned outside of conscious deliberative acts.

Karen Jones has argued that it would simply be a surprising thing to find out that we should by default trust testimony when the phenomenon of testimony is so thoroughly diverse.[179] Testimony can cover a wide range of subjects, be given by a wide range of people to a wide range of audiences, and occur within quite different social conditions and institutional settings. That in any case the same approach should apply may strike some as not being appropriately sensitive to changing testimonial conditions. As a result, there are at least some concerns raisable in applying so universally the antireductionist principle (i.e. default trust) to all testimonial contexts and content. Indeed, Faulker has stated that one of the most difficult things to stomach about antireductionism is the ease with which it seems to simply ignore issues of epistemic responsibility. Surely we *ought* to critically monitor testimonial content and sources.

One can see from this overview that there is some division on the nature of testimonial justification. The discussion that follows will treat testimony with a modest lean towards reductionism in regards to the project under examination here. The reason for this is that it strikes me as intuitively plausible and also well suited to the epistemic ends this thesis pursues. I will be careful however to leave open a door to those of antireductionist persuasions, or to a more sophisticated synthetic account of the sort Graham entertains.

4.4 Positive Conditions in Belief and the Reductionist Model

The independent reasons of an expert for believing some proposition *p* will be quite different from our reasons for believing *p* through them. It could be that they believe it for highly technical experimental or theoretical reasons, whereas reductively we would

[177] Ibid, pp. 94-95.
[178] R. A. Naulty and P. J. Sheehan (1975).
[179] Karen Jones (1999).

come to believe *p* through them by having observed on previous occasions how their expertise and claims of that type have related to truth, and then accepting on that basis other claims that they make (*p*). As a basic first condition of testimonial acceptance it seems reasonable to say that we require more than this however. We require reasons to think not only that a given proposition is true, but that it is justified in being held as true by the experts. That is, we require reasons to think that the belief in question was adopted on justifiable grounds by those asserting it, aside from what reasons *we have ourselves* for adopting that belief in testimony.[180]

Now, immediately we have run into an obvious problem, in that most experts will acquire much of the knowledge of their expertise themselves only through testimony. A doctor does not learn the greatest proportion of his profession by discovering medical facts himself, but by relying upon a huge epistemic network of experts before him. For the purposes of discussion then I will speak as though the justification of *some* experts in believing propositions non-testimonially can become the justification of experts within that expertise as a whole. I will speak as though experts who discover (and justifiably believe) facts in their domain confer that justification to the expertise itself in relation to those facts. This is for efficiency, but it should be understood that no deeper epistemic implications need to be drawn by this way of speaking.[181]

The belief in the justification in the natural setting can be posited as part of what constitutes the best explanation for the success of the beliefs in question, and in relation to the prior improbability of their being true. It would be an amazing coincidence, for instance, if the many scientists and engineers employed in successfully landing a person on the moon were simply *unjustified* in holding as true the manifold beliefs they utilized in doing so.[182] Any achievement of that calibre will require massive, multi-layered reliance upon background knowledge that draws across multiple disciplines, and requires thousands of steps to be undertaken successively and carefully; the failure of any one of which could render a successful outcome impossible (sometimes, spectacularly and tragically). It is true that this success *could* have been chanced upon without good reasons being held in relation to the propositions used, but this simply asks too much to be seriously entertained. It would make each moon landing a miracle of blind luck.

It is similarly so - but less dramatically - in the case a doctor determining which drug from among tens of thousands available to him should be administered to treat a patient. The likelihood of any particular drug being correctly suited to treat a particular disease is low, and so the success of a drug in treating that disease points us towards the expertise's

[180] Faulker agrees with the relevance of the justification of testifiers to hearers in asserting the propositions that they do, noting also that their intentions are relevant. He notes these points as having no parallel to other primary ways of knowing, such as perception. See Faulker's (2002).

[181] What is raised by this manner of speaking is questions about whether abstract entities like 'domains' and 'communities' can 'know' things, and be 'justified'. As I say, I do not commit myself to a view on this; I speak in this way for the purposes of discussion here, simply because it is useful in order to avoid tedious articulation of the fact that only various segments of a domain (many times, *historical* segments) acquire the relevant propositions directly. For discussion of the application of epistemic terms to abstract social entities, see John Hardwig (1985).

[182] Examples borrowed from Alvin Goldman (2001)

relevant justification in that case and others.[183] How so? It gives us reason to think that the beliefs held by the expertise are justified, *and in a way that relates to the nature of the expertise itself* (as a best explanation for its success).

Consider a case in which expertise E contributed the set of propositions Q in the bringing about of achievement A, and achievement A is highly improbable by mere chance alone as assessed in relation to our background knowledge. Part of the best explanation for A's obtaining, given its improbability, and E's implied success by this, is that Q is (at least approximately) true and E was justified in believing Q. From this inference we determine something about E and the veridicality of the principles of inquiry employed there by E's ability to reliably produce significant truths. And from inferences to the probable justification of experts in believing those propositions that ground our trust in their reliability, we can infer to the expertise's most probable justification in regards to future claims, on the condition that they are of the same type and are recognised as stemming from the same conditions of expertise that yielded the formerly successful ones. In other words, from the observed remarkable successes of an expertise we can infer to the justified status of the beliefs involved with that expertise as part of the best explanation for their being apparently correct, and from this we can then reasonably infer to the probable justified status of future claims. So while at time *t* we may not know anything specific about the nature of the evidence that is drawn in support of Q, or enough to enable us to independently believe Q based upon it, we can nonetheless infer something about the nature of the evidence supporting Q (as seen by E) by observing successes involving Q.[184] This inference is modified by Q's improbability. The more unlikely a proposition (or set of propositions) is to be true, the more reason there is to think them justifiably held in the case that there are good independent reasons for thinking them as being true. In this way, our confidence in a domain's justification scales according to how improbable the propositions are, in light of their apparent truth. Accordingly, stunning successes (the Manhattan project, the human genome project, the Apollo space missions, etc.) can provide us substantial confidence in the domains of inquiry we understand to have been involved, both in terms of their ability to provide truths (as evidenced by the stunning success), and in the principles and methods of their inquiry (so far as we have knowledge concerning that). Since many achievements draw upon numerous domains of expertise, resounding successes can cast a wide net of measured validation. Moreover, all domains of expertise utilising similarly recognised principles of inquiry are marked by these successes in a *limited* way, even if those expertises are not directly involved in those marked successes themselves.[185] This 'trickle-down' effect is particularly salient in light of the apparent difficulty of ascertaining the veridical connections of expertise.

[183] Putting aside obvious problems in that case, such as that the disease just ran its course and the drug in fact provided no benefit in aiding to cure them.

[184] These dramatic tests are not 'litmus' tests; we would not want to say by any means that a *failure* of the Apollo moon missions meant that the propositions contributed and relied upon by any given expertise were *false* (or unjustified). Clearly the overall project could have failed for any one of tens of thousands of reasons independent of that particular contribution.

[185] I.e. the human genome project and similar successes were all 'victories' as much for the scientific method as they were for the particular fields directly and indirectly responsible for them.

The picture I am presenting, then, is one that draws emphasis outside of simple linear, one-to-one correspondences between expertise and truth. The web of affirmation and trust is viewed instead as often being a massive network of innumerable successes of variable magnitudes, including many notable and profound ones, and of interrelated, interdisciplinary reliance's and confidence-sharing.

To summarise: the constructed reductionist position is that the correlation between truth and testimony is evidence for the probable truth value of a new proposition, of the same type and from the same source, so far as the new proposition was justifiably formed from roughly the same conditions of expertise and inquiry which yielded the truthful propositions formerly observed. So far as this is the case it is justification for the further acceptance of testimony. The observation of past veridicality does not constitute 'direct' evidence, as it were, for the truth of any new proposition, but instead acts as indirect evidence that increases the probable truth value of new propositions inferentially. A limiting condition on the acceptance of testimony is that the expertise in question is believed to be most probably justified in holding the claims asserted. This can be inferentially established as part of a best explanation for the apparent truth of the claims already made, and can then be inferred for new instances of testimony insofar as those propositions are understood to proceed from the same conditions of expertise.

If this picture is correct then our beliefs play a definitive positive function in supporting testimonial practices. Observations made between expertise and truth are important for more reasons than just reductively establishing *that* an expertise has a veridical connection, but also mutually providing the basis from which it can be inferred that the expertise is justified within itself in making and holding the claims that it does. Of course it could be denied that our belief in the initial justification of propositions is required, but such a prerequisite seems to be intuitively powerful and, moreover, normatively compelling. It seems just flatly reasonable to say that we epistemically *ought* to believe that others are justified in their beliefs before we accept them as our own.

4.5 A General Delineation in Testimonial Deference

The implications of epistemic deference are, as well noted by now, substantial. A reductionist model of the sort outlined above is well placed to limit the number of ways in which error can be introduced through testimony, since we are required to actively engage in testimony by having positive reasons grounding our decision to defer to others. The justification involved in deference is earned through our connecting a source of expertise and testimony type to observations regarding its truth status. If this view of testimony is correct then it is clear that we rely extensively upon checks and balances in order to gauge the reliability of sources of testimony. I want to make a further claim however, which is that testimonial practice should be centred on our most unambiguous and reliable forms of inquiry, in which the truth and falsifiability of propositions is most easily available. The normative claim is that we should seek to place (loose or tight) limits upon our deferential practices in such a way as to reduce the possibility of undesirable epistemic, veridical, and prudential consequences resulting. I suggest that this

is best done by linking deference to domains of inquiry that most rigorously and openly evidence their reliability, and allow for (and often force) correction and revision.

We can express this claim by way of two epistemic principles, intended as applicable to cognitively normal people possessing a sufficiently mature level of knowledge:[186]

> Weak Deference Claim: We should consciously proceed with measured caution before deferentially accepting primary propositions from sources of epistemic authority, where the propositions have no firm empirical character, or where the propositions advanced otherwise lack substantive ways to determine truth and falsehood.
>
> Strong Deference Claim: We should provisionally limit ourselves from accepting primary propositions from sources of epistemic authority, where the propositions have no firm empirical character, or where the propositions otherwise lack substantive ways to determine truth and falsehood.

Both principles seek to (weakly or strongly) limit testimony in relation to primary propositions. Primary propositions are direct statements, such as of the familiar type 'god exists' or 'it is wrong to kill others for the pleasure of it'. Secondary propositions are those such as 'most people believe that god exists', or 'most people think it is wrong to kill other people for the pleasure of it'.[187] Both principles emphasise a level of individual autonomy in social epistemic relations and endorse a critical attitude. We are explicitly called upon to weigh up the benefits and risks involved in deferring outside of our most open, rigorous, and confidence providing areas of knowledge. The weaker principle is more liberal by allowing for deferential moves to be made, but only after some level of reflective consideration. The stronger principle implies that any exceptions we would want to make are worked out (either individually, or socially) prior to testimonial settings.

Underlying both principles is a *responsive deference* claim, under which our approach to testimony is taken to be sensitive not simply to the consequential implications of the propositions involved, but also to the way in which the propositions can be found as either true or false. That is, the more a proposition makes itself available to being shown false, and the more it resides within areas where truth can be determined with a high degree of confidence, the better it is, and the more appropriate it is as a candidate for deference to others on.

Both of these points need defending. In relation to the first, I take it as a fact about us that the way in which we meet testimony is modified by the consequential implications of acceptance. That is, our demands for confidence in authority are modified by what we see as being the direct or indirect (veridical, theoretical, social, etc.) significance of accepting as true some proposition. In this we can be expected to bring higher levels of scrutiny and demands for confidence in relation to propositions whose consequences we understand to

[186] We would probably want to make exceptions for children, for instance.
[187] This distinction is owed to Goldman (1999), p. 352.

be important in some way for ourselves or for others. It is to be understood that both principles outlined before apply in conjunction with this feature. As an aside, it is noteworthy that this is open to a certain kind of abuse. If for instance a group of people *A* have a longstanding cultural interest in the truth of a certain belief *p*, and an otherwise trusted epistemic authority claims ~*p*, it is possible for *A* to demand such overwhelming evidence for the truth of the claim ~*p*, given the massive value and the central role that *p* has within *A*'s way of thinking, that the demands made can essentially never be met. If for example the problem of induction is invoked in order to avoid some alleged and apparently unpleasant or undesirable scientific truth, I think it would be reasonable to say that something has gone wrong in relation to the evidential standards being required. The epistemic concern in such a case is that since science operates essentially inductively, that it will never be able to interrupt beliefs of the highest importance in the case that one is sceptical about induction. But this is *in effect* tantamount to taking a position of faith, since no body of evidence will ever suffice to meet the required levels to overturn a sufficiently valued belief. This, I think, must surely constitute an abuse of the principle in question, whose purpose is not to make inviolable certain beliefs but justify high epistemic standards in relation to matters of consequence.

The second point is that our willingness to defer to others should be modified by the character of the propositions. Much hinges here upon what 'firm empirical character' and 'substantive' conditions of truth and falsehood amount to. I consider a proposition to have a firm empirical character insofar as it is possible to make observations, in conjunction with other propositions, which indicate in a reasonably strong way its truth or falsehood. A scientific theory makes claims about the way the world is, from which predictions can be made that when tested allow for either the corroboration or the falsification of the theory. A lot of reasoning however does not end in the same kind of discrete observational entailments. It is true of course that some forms of reasoning are quite stringent, such as in the case of deductive logic, but deductive derivations make up only a small segment of our overall reasoning activity.

Many will recognise the picture I have painted of empirical inquiry as overly simplified, and it is true that there are a large number of epistemological nuances that complicate matters. For example, one consequence of the theory-ladenness of observation is that there can be no fully decisive observational falsification or confirmation of advanced scientific theories. Moreover, as Duhem showed, science is not like mathematics: from the falsity of one of two rival theories we cannot deduce the truth of the other (for example, we cannot deduce from the falsehood of the particle theory of light that the wave theory of light is true, or visa versa).[188] In this respect there can be no truly 'crucial experiments' in science. But such difficulties placed to one side, I nonetheless consider it to be rather uncontroversial that empirical inquiry carries a distinct epistemic advantage insofar as truth or falsehood therein turns upon meeting or failing to meet conditions of empirical observation, over and above the satisfaction of merely logical conditions. The world is by nature simply more refutational than are the bounds of mere reason alone, but far from being a weakness of this kind of inquiry it is a decisive epistemic strength. We can *comparatively* know more clearly in many cases when we have it right and when we

[188] Pierre Duhem (1954).

have it wrong in science, for the conditions that show empirical propositions as true or false are more numerous and open to more intensive examination. It is this epistemic qualitative situation that I take to amount to 'substantive' conditions of truth and falsity.

Deferential preference cannot be limited to only empirical forms of inquiry however. What, after all, is the nature of mathematical or legal testimony? Neither seems to fit an empirical category, as each is concerned with either abstract and socially constructed facts, but both can offer a high degree of verifiability, falsification, openness, and concreteness, and in this they do not seem to be the appropriate sorts of things to be deferentially advised against (if anything is to be). Of course – and I stress the point – closer attention needs to be paid to the specific nature of a proposition received is in any case. Whether a law '*x*' exists may be a much more concrete matter than whether law *x* applies to situation/act *y*. In the latter case, and if it is of a particularly vague or uncharted case, much may rest upon the interpretation of the law, and this may vary from one individual to another, and require vast amounts of legal experience to determine. In such cases propositions of the sort described (*x* applies to *y*, *x* does not apply to *y*) may fail to meet unprejudiced deferential standards. And indeed it is true that propositions of any type may have variable degrees of empirical engagement and conditions for the observation of their truth, and so require that our approach to propositions from those areas and types is matched in accordance with that detachment. For this reason it is clear that the success of these normative principles relies extensively upon our level of background knowledge regarding inquiry. Our understanding of the world and of forms of inquiry is of central importance to how we understand the deferential status of particular claims.

I am willing then to allow that others may take a different view in regards to what kinds of things or conditions satisfy the principles. That is to say that someone else may have different ideas about what constitutes a propositions meeting of the condition of having a 'firm empirical character', or what otherwise constitutes sufficiently strong truth conditions. My underlying goal is not to provide any hard and fast, discursively neat delineation of acceptable and doubtful testimony, but rather to outline a reasonable approach to testimony that develops and takes account of a certain critical attitude that aims to be responsive to the type of testimony we receive. If only the most concrete answers to difficult epistemic questions of testimony will do, then what has been advocated here does not achieve what is required of it. But we ought not to consider such high demands as being necessary to improve upon our epistemic condition. As mentioned, the greater utility of these principles may just be in their lending towards the development of a conscious attitude that urges tentativeness before deference to testimony involving primary propositions of certain classes/types.

There should be no fears harboured that this is logical positivism in another inflection. These principles offer nothing in relation to questions about what it is possible for any individual to know, alone, rather than (and only) what they are advised to do in the face of testimony, and so what they might know *through others*. The principles given aim to reduce epistemic surplus by placing certain violable restrictions against the types of propositions we find put to us. They thus aim to strike a difficult balance between

encouraging doxastic autonomy and responsibility while maintaining deferential pathways towards epistemic authorities.

To give late attention to the antireductionists among us, I think it would be possible to incorporate these general principles into a list of potential defeaters of testimony in any case. So for instance, our knowledge that a speaker is upholding primary propositions which do not relate to the way the observable world is, or cannot provide similarly stringent truth-status conditions, is a defeating condition of deferential acceptance. Again, this would not make any claims about what it is possible for any individual to know, or about what it is meaningful for them to express, but about what it is possible for any of us to know through others.[189]

4.6 Against Borrowed Moral Knowledge

The kinds of primary propositions that we are guided away from in relation to deferential practices by the previously articulated principles are going to be most typically those of an ethical or metaphysical nature. While many would in any case view with some scepticism the idea of borrowed moral knowledge, Karen Jones has argued that there is a significant place in social practices for moral deference to others.[190] Jones's argument rests upon two claims: that a) people can be morally 'blind sighted', and that b) there is no clear-cut, hard and fast distinction between testimony and argument.

The truth of the first claim is easy to establish to our satisfaction. Experiences can shape our moral understandings, and as a result of people having diverging experiences they can come to have different degrees of 'moral perceptiveness'. We can readily imagine the case of a sheltered prince whose detached and privileged life, in conjunction with the socially prevailing hierarchical class beliefs of his culture, have left him serenely ignorant of the moral and social situation of the palace slaves he sits a crest of; a condition of which he might stand a good chance of being remedied in the case that it were possible to transplant him into their position for some extended period of time. The prince, we can say, simply fails to 'see' the unethical situation of the slaves as such, in part because he has not been attuned by his own experiences to be morally perceptive in that respect. It follows as a result of people's having unique experiences that some may be in a better position than others to recognise the moral qualities of certain situations and acts. It is on the basis of this connection between experience and moral vision that Jones makes the normative claim that "those who lack the relevant perceptual skills in a given moral domain must rely on those who have sharper moral vision".[191]

The second claim, if true, would blur the distinction between the moral testimony and moral argument. The possibility it points towards is that even in cases where we thought

[189] A certain paradox arises from this. If we are to allow for children to evade the application of these principles, then it seems on one interpretation that children *can* know certain things through testimony that adults cannot.
[190] Karen Jones (1999), p. 56.
[191] Ibid, p. 63.

we were responding to argument, we could have been simply responding to testimony. Jones illustrates this point analogically: expert witnesses in trials may do as much to testify as to argue a point. They may offer arguments in support of either the defence or the prosecution at trial, and try to explain their findings to the jury as much as is possible, but oftentimes the jury will simply not understand *enough* of what is necessary to accept the argumentation on its merits alone. They may instead supplement their understanding of the argumentation presented with the implicit or explicit testimony on the part of the experts that the evidence presented constitutes something appropriate or sound in the way of argumentation – something they could not establish themselves without more extensive relevant background knowledge. For instance, a jury may not understand why a certain percentage on a blood sample test constitutes a very high match between two people (and believe the expert when she testifies that this is the case), but understand very well the basic argument as to why blood samples that relate to one another is evidentially significant for the guilt or innocence of those on trial.

On one hand, it is a simple truism that argument and testimony are intertwined with one another, for it is a basic fact that any argument to the truth whatsoever is at the same time implicit testimony to that effect. One cannot consistently present an argument as true without simultaneously implicitly or explicitly endorsing it as true. But that testimony is always found alongside argument does not mean that the argumentation alone cannot be, or is not in fact, the singular force in a person's coming to accept it or not. Jones's point is a finer one however; she wants us to accept that even in cases where we think we are accepting an argument on its independent merits, we could be just as much relying on testimony in order to do so.

One response to this could be to claim that while the jury situation might have the *appearance* of being partly argumentative, that it is *essentially* or *basically* testimonial. The jury is engaged to understand as much as they can, but as all the processes and background assumptions used to reach the expert's judgment cannot be explained to them in any efficacious fashion, what the epistemic situation in fact amounts to is basically testimonial. The reasoning the jury can understand establishes (or aids to establish) their confidence in the expert's competence and processes, and moves them to accept the experts testimony on that basis, rather than by their seeing the full light and shape of what is argued. This does not blur the distinction between testimony and argument, but expands the number of cases of testimony to situations in which we thought that it was argument at work. This however could imply that in so many of the cases in which we thought we were responding to argument from others in every day life, that in fact we were just deferring (unbeknown to us). This may be more than what we are willing to allow for. In any case, the more important point for us to consider here is a question of what this is all supposed to normatively allow for? That testimony is involved in some (or many?) cases of our accepting argumentation is a side issue to whether we normatively ought to rely *wholly* upon the testimony of others for specific moral claims. This point can be accentuated further by noting that Jones's example (of the jury) involves specifically *non-moral* argumentation. If argument and testimony intertwine only in this way in non-moral domains, what normative implications this might have for *moral* testimony remains unclear. In the case that we do in fact make use of moral testimony,

this does not in itself mean that we *ought* to (and it is a normative claim that is the essence of Jones's argument).

The centrepiece example used by Jones to support her case for borrowed moral knowledge is the presumably hypothetical case of a man named Peter living in a cooperative house and facing a dilemma over membership decisions taken by members there. Candidates to the house are subjected to interviews in order to ascertain their suitability, with any of the members living there holding the power to veto their successful application. After some time Peter came to find the decisions taken in this process intolerable. Three men who had applied to live in the house had been rejected on the basis of what had been perceived by three white and coloured women as sexism (and, in one instance, racism). Peter himself was not sexist, and was sincerely committed to the idea that sexism was something bad and to be avoided, but he could not see for himself in any case what reasons the women had for judging the men guilty of it. When Peter pursued an explanation it was put to him that it was not in anything the male candidates said, but rather in what they did. He was not aware of anything in their behaviour that indicated sexism, but in inquiring further he was told that it was not because the men had stared at the women's breasts, or anything overt of this sort, but simply because of the way the men had made them feel during the interview process. They described this feeling as resulting from a kind of "non-presence" during questioning, which resulted in them feeling unimportant. This, they judged, was evidence for a sexist attitude. As a result of his need to feel that he endorsed the membership decisions made in the co-operative and his inability to find agreement with the decisions that the women had made, which he judged as arbitrary and discriminatory, Peter chose to leave.

Peter's case is of the simplest type, since it does not involve his accepting whole new moral principles previously foreign to him, but applying already held moral principles against sexism to certain specific cases.[192] Jones is confident that in such a case, Peter, while showing particular notable virtue in his leaving (for it took character to require that the decisions made honestly reflected his own position), ought to have simply accepted the women's judgments that the candidates were guilty of displaying sexist biases.

Jones is sensitive to the dangers inherent in moral deference, and particularly the concerns people have over them. As a result she tries to alleviate doubt by making the distinction between the practice of moral deference and the principle of it. That is, we might disagree about whom to trust and how they come by that trust, but these issues she thinks "do not arise when considering the question of what sorts of circumstances, if any, are ones in which an agent should rely on someone else's moral judgment".[193] This is questionable, for it seems that it is *precisely by* the practical difficulties and concerns involved with the idea of a deferrable moral expertise, and how we would recognise it, that few if any people appear to be proper objects of moral deferral. In the minds of many, those difficulties make it quite unlikely that circumstances could be arrived at in which moral deference is allowed for or commendable. So substantial, in other words, are

[192] A third possibility involves the relative rankings of various moral principles in relation to one another in importance, which we could possible defer to others on.
[193] Ibid, p. 64

issues over the implementation of the idea of moral deference that it cannot usefully be separated from questions of the very idea of it.

On what conditions, then, is it said to be advisable to defer to others on moral issues? Jones argues that the basic default stance towards those claiming moral authority ought to be one of distrust. The conditions she lists as changing the stance to one of trust include knowledge of the person's character, knowledge that they are self-critical, knowledge of any hidden-agendas they may have, and knowledge of whether they have the sorts of experiences necessary to provide the basis for the kinds of moral claims they make. She comments as an aside that those who are possibly most worthy of our morally deferring towards are those that are least inclined to offer moral testimony.

Questions over how unique moral expertise can be identified, of how the value of moral deference in those cases can be recognised by those unable to recognise the moral truths posited, and how the justified concerns people have over the moral error possible in testimony can be ameliorated, are central to the question of the viability of the idea of moral deference. The moral truth we would have to obtain from others would have to be judged to supersede in its value the dangers involved in engaging in moral deference in the first place, but as a result of our being unable to recognise it as moral truth our ability to recognise its value as such would be greatly restricted.

We may question how it is that we can know that what is put forward in moral testimony is in fact a genuine moral truth. Just as Popper's Marxist, after having adopted a Marxist way of 'seeing' the world, immediately observed many more immoralities and violations of ethical principles around him than before, and came to see them in even the smallest and most minute of things, there is no way to be sure in any case that what we morally perceive in even our own case is genuine, and not simply the phantoms and aberrations of our adopted ideology, unique history, or interpretative stance. We might say, for instance, that from our personal experiences in having worked on a factory floor that it is clear that there are many things deeply ethically wrong with the capitalist mode of production. But this moral determination could just as easily be a fabrication of ideology as it could be a genuine moral insight backed – and indeed made possible by – ones life experiences. As in our case, so to in that of others; there is no way to be sure that the claims of others are the pure results of their unbiased experiences and properly functioning moral intuitions, perceptions, and reasoning. This may be deeply unfortunate, because there is nothing to say that there *aren't* many unethical states of affairs visible only to a morally-acute few. But the solution to this unfortunate condition would not seem to wisely be made out to be our moral dependence upon others.

How could we even arrive at a position where the reliable ethical capabilities of others could be inferred? One might argue that it could be possible to establish sufficient moral confidence in the judgments of others in the case that it were possible to establish a sufficient track record of a person's moral insights being correct, insofar as the truth of those moral insights *is* available to us. While this may be possible in principle, given that our requisite confidence in others in regards to their moral judgments would have to be very high, in practical terms there is not likely to be at any time, anywhere, any group of

people whose moral insights are found to be so uniformly correct and novel that they can serve as a basis for our deferral towards them on truth-unseen moral principles or applications. That is, given that issues of morality are of such crucial importance to us, in order to acquiesce to others on ethical matters our confidence in them would have to be substantial and well established. Such well founded and concrete confidence as this, while in principle possible (as so many things are), is simply never likely to arise. Insofar as current world social practices allow not simply for moral advice giving, in which we gain access to the moral reasoning of others, but moral deference to authorities, there are reasons to be concerned and epistemically dissatisfied.

Jones admits that there is still much left to be done in working out the details of what conditions satisfy those that normatively compel us to defer our moral judgment to others, but those of us who are - as Popper put it - "alive to the dangers inherent in all forms of power and authority", and who recognise the quality of our moral beliefs as often being of immense importance, will need especially powerful reasons to overcome their doubts.[194] These reasons, high as they are, are not supplied in her defence.[195]

[194] Karl Popper (1963a), preface to the 1st edition.
[195] For other doubts regarding moral testimony, see Philip Nickel (2001). Admittedly, Nickel does not view the objections he makes as applying to the kind of moral deference Jones advocates.

A Critical Project

5.1 Dealing with Irrationality

In light of an increasingly detailed cognitive reality about us, we well may ask what it is that we are to do, individually and collectively. What can we do to better our own epistemic position and that of those around us? We have seen that biases elicit their effects widely, and we can suppose there to be little guarantee of a total absence of bias or error in any case. How then should we proceed?

The reality of the question is rather more complicated than this however, for in even framing it in this way we have assumed that there is some agreement that there is a problem here at all. We have assumed that others along with us have shared interests in improving our epistemic condition along rational lines. But such interests are far from universal; some would claim no affiliations with or affections for the kind of rationalism implicitly underpinning the problem as outlined. What then can be said or done about those who would claim no interest in being 'rational' or seeking 'truth'?

Popper, like others, was acutely aware of the limitations of rationalism in this respect. Rationalism, as considered, is not the view that reason can provide many truths about the world *a priori*, but rather a certain fundamental attitude of placing substantial value on argument and evidence in determining what it is that we believe. It is a willingness to admit the possibility of error in ones thinking, and the error of others similarly; to be responsive to argument and to engage with others in mutual critical growth. As such, it is a view which denies that tradition carries any special weight of authority on matters, and forces us to recognise that we are indebted to others for so many of our beliefs and so much of our acuity in reasoning. At the heart of these rationalist principles is a certain traditional conception of truth, and the belief that truth of this kind is possible and available to us. Yet rationalists must face a certain unavoidable truth that "no rational argument will have an effect on a man who does not want to adopt a rational attitude".[196] There are as such profound limitations to the purchase of rational argument, and it ends at the will – at the decision within us to be rational (or not). No argument can establish the cause of rationalism for those who are sufficiently resisting or ignoring of it.

Facing realities of this kind, what extra-rational grounds can be appealed to in order to lend aid to rationalism's cause in the eyes of others? The obvious choice for many is to seek to do so on moral grounds, and this is precisely what Popper attempted. He encouraged a broad 'faith' in rationalism precisely because he considered it to be of great moral importance that we adhere to rationalist principles. Irrationalism, as he understood it, has a long and sordid history. It has many manifestations, but Popper saw it as most commonly dividing the 'rational unity of mankind' in two: between those who are responsive to argument and evidence, and those who can only be appealed to by emotion and passion. The problem was not that irrationalism did not employ argument, but rather

[196] Karl Popper (1962), vol. 2, p. 257

that it feels no obligation towards its use in any consistent way, and thus uses it and discards it as it pleases.

Identifying irrationalist aspects in thinkers such as Plato, Hegel, Whitehead, Toynbee, and more generally in the fashionable intellectual trends of his day, Popper hypothesised that such sentiments benefited from conditions of ignorance. He drew an analogy between this and the case of a chaplain in a play of Shaw's who lusted after Joan of Arc's execution only in the absence of immediate knowledge of what that entailed. Upon seeing her at the stake, he cried out "I meant no harm. I did not know what I was doing… If I had known, I would have torn her from their hands. You don't know. You haven't seen: it is so easy to talk when you don't know. You madden yourself with words… But when it is brought home to you; when you see the thing you have done; when it is blinding your eyes, stifling your nostrils, tearing your heart, then - then - O God, take away this sight from me!".[197] What is described in this instance is a species of availability error wherein we may find ourselves approving of something so long as it remains in the abstract, such as is contained in a mere idea (e.g. the punishment of people as a matter of justice and retribution), yet be forced to position ourselves quite differently when confronted by its entailments in a far richer and much closer proximity. The chaplain in question failed to imaginatively draw a connection between his abstract views and what they demanded in order to be meted out. Similarly, argued Popper, many irrationalists fail to appreciate the full substance of what their intellectual beliefs entail in relation to the moral realm. And so by framing the question of rationality in a moral light, rationalists may hope to engender rationalist commitments by directing people towards the moral dangers present within that broad intellectual standpoint.

Popper's point may be strengthened somewhat if we allow that some of the motivation for irrationalism not only survives within a certain kind of moral ignorance, but originates from a basically ethical place. I suspect that some or even much of the motivation towards scepticism about truth, argument, and evidential standards, derives from concerns over the moral implications of such concepts in their application. Contrary to the enlightenment conception of truth, which saw truth as essentially liberatory and freedom enhancing ("man can know, thus he can be free"),[198] some have viewed the idea of truth very differently. From an anti-enlightenment viewpoint, truth can be seen as a tool of the powerful for the suppression and domination of the weak. It is something invented and used by people in order to impose ideas upon others and to contain them in their place. Truth in this way is essentially limiting, confining, diminishing, negative, and wielded over people in the special interests of others. It is thus something to oppose for obvious ethical reasons. If this way of conceptualising about truth is what causes much of the enmity towards rationality, then an argument that appeals to the moral inclinations of people can hope to challenge them in such a way as to invert the judgments against truth.

While Popper's argument has some merit, it is unclear how far it can carry us. If it is the case that moral concerns drive some irrationalists in a rejection of reason and evidence, then drawing attention to and re-emphasising the moral chasm possible in irrationalism

[197] Ibid, p. 257, quoting from George Bernard Shaw's 'Saint Joan'.
[198] Popper's summation of the enlightenment motto. Karl Popper (1963a), p. 7

could be an effective method towards persuading some sceptics to reevaluate rationalist principles. However, it must also be acknowledged that the scepticism of some that calls doubt upon rationalist principles and concepts may just as easily call doubt upon many widely held or standard conceptions of *morality*, or even on the whole notion of morality itself (in which case, it would *not be* the same kind of scepticism about truth that is driven by widely shared ethical values and attitudes). In such a case, the Popperian method outlined would fail to achieve what it seeks. This is not an indictment against Popper's argument as such, but a restatement of the fact that for those who seriously doubt traditional conceptions of *both* truth and morality, there may be little that can be convincingly said.

While Popper appears to have been focused upon the ethical dangers of foregoing rational principles in a broad sense, others have emphasised a connection between belief and ethics in a much closer way. In his best-selling polemic against religion 'The End of Faith', Sam Harris urges readers to recognise that "belief is not a private matter; it has never been merely private. In fact, beliefs are scarcely more private than actions are, for every belief is a fount of action *in potentia*".[199] With an eye to the way in which various religious beliefs have caused undesirable consequences, he stresses the point that beliefs define our vision of the world, determine our emotional responses, and dictate our behaviour.[200]

The eradication of the private sphere of belief in search of measures to enforce rational considerations has historical precedent in the writings of William K. Clifford. In 'The Ethics of Belief' Clifford takes a distinctly deontic stance in regards to matters of belief. Our believing something, he holds, is right or wrong not in relation to whether it turns out to be true or false in matters of some moral consequence, but whether the belief was come into in sound ways or not - quite independently of what consequences it brings to bear.[201] The moral quality of any belief is given in virtue of whether it is established reasonably or not – not at all by whether it (in fact) causes harm or generates good.

Clifford asks us at the outset to imagine the case of an old and worn passenger ship that was to be put to sea. In consideration of whether the ship was seaworthy, the owner questioned to himself whether he should rescind the orders for its planned voyage and have it overhauled instead. Allaying these doubts about its condition, he told himself that the ship had made many successful voyages and had survived many sea storms previously. It was not to be suspected, he thought, that this time was to be any different from the last. And with so many lives onboard surely providence would see her safely to her destination. As circumstances would see it however, the ship would not survive the storm that hit her mid-voyage, and she took all on board down to the crushing oceanic depths below. Should we consider the owner as bearing any moral responsibility for his beliefs and resulting choices, given that he convinced himself the voyage would be fine, removed from himself his fears and doubts, and let the ship and its passengers sail to what would be, in fact, their end? In Clifford's view, the owner was plainly responsible

[199] Sam Harris (2004), p. 44.
[200] Ibid, p. 12
[201] William K. Clifford (1877).

for the deaths of those onboard, for the reason that given the evidence before him in this hypothetical scenario it was inappropriate for him to have convinced himself that the ship would sail without issue. But Clifford goes further than this, and holds that *even if* the ship *had* made its crossing successfully, or even if the ship had not been in fact an unworthy ship at all (for all that anyone knew, or could have known), he still would have been wrong – that is, *morally* wrong – in holding the beliefs that he arrived at. The only difference in such cases would have been that the ships success in crossing would have concealed the man's error. Clifford thus frames the issue of belief squarely within a deontological conceptual and linguistic framework. For him, the broader question placed before us is whether we have a *right* to believe what it is that we do in any given case. We can accidentally be right and accidentally be wrong in matters of consequence, but the heart of any moral issue always rests upon whether the evidence before us could have most reasonably justified our beliefs, and it is at that point that moral guilt or innocence is determined.

Clifford rightly recognised (more will be said on this later, against Clifford's position) that many beliefs do not lead directly to actions. Beliefs can instead be incorporated into a broader background body of belief which gives guidance to future action. If a certain belief promotes happiness for instance, such as a belief in a person's continued existence after death, this will indirectly influence future actions by influencing the mood and attitude of the person, given basic psychological facts about us. Clifford suggests in addition however that the impacts of beliefs of this kind when ill-held can be even more greatly extended than this, as any irrationally supported belief will (he says) set a dangerous irrational precedent that may cause more of the same. And so "no real belief, however trifling and fragmentary it may seem, is ever truly insignificant; it prepares us to receive more of its like, confirms those which resembled it before, and weakens others; and so gradually it lays a stealthy train in our inmost thoughts, which may someday explode into overt action, and leave its stamp upon our character for ever."[202] There can as such be no completely 'innocent' beliefs in Clifford's estimation, since even those ones which do not themselves directly result in actions ineluctably indirectly contribute in various ways towards others. And if these beliefs are acquired in unreasonable ways, then moral guilt is incurred by the believer.

It is by way of this reasoning that the eradication of a private sphere of belief is set about. If all beliefs directly or indirectly influence action, and the judgment of actions with respect to one another something which falls largely within the moral realm, then it can be argued that each and every belief an individual possesses is in the public interest. There can simply be nothing like a private or innocent belief – beliefs that people are 'entitled' to by way of their own rights to self-thought - because there is no belief that does not lead directly to action, indirectly guide it, hold the future possibility of causing action, or in itself not set a precedent in its irrationality towards future irrational belief. He thus writes that "belief, that sacred faculty which prompts the decisions of our will… is ours not for ourselves *but for humanity*".[203] He encapsulates his position famously as being that "it is wrong always, everywhere, and for anyone, to believe anything upon

[202] Ibid.
[203] Ibid, emphasis mine.

insufficient evidence."[204] In this way moral duties become epistemic duties, and we are, at the penalty of wrong doing, absolutely ordained to think rationally.

The position presented is a radical, intrusive, and even totalitarian one. It suggests not that abandoning rationalism leads to a greater risk of doing wrong or harm, but that one *immediately does wrong* when ones thoughts are not properly rationally formed. In contrast to Popper, who acknowledged that some irrational beliefs are quite harmless, Clifford's account accuses that there is harm done in any less-than-rational evaluation or belief-forming process of thought. This is precisely the sort of thing we can imagine inspiring the aforementioned anxiety and distrust of some in the calls of others to 'truth' and 'reason'.

5.2 The Belief-Laden Nature of Action

Beyond Clifford's masterful command of prose and phrase, what can be said of the philosophical position outlined? It appears to be problematic in a number of ways, some of which cut more deeply than others. A first issue results from the deontic framework on which it stands. We may question how exactly we are supposed to go about thinking of an ethics of belief divorced from considerations of consequences, and query what such an ethics could possibly amount to. We may object to its counter-intuitive status: for Clifford, the rightness or wrongness of a belief is independent of the results it procures, and so believing good-causing irrationalities is as *morally* condemnable as believing wrong or harm-causing irrationalities. To those not accustomed to approaching ethics from a deontic stand-point this statement is likely to appear as quite intuitively awkward. Surely, we might reasonably think, the *moral* nature of belief is best determined not by its *rational* character but by its moral consequences. We may even suspect that it is indeed the consequences of irrational believing that drive Clifford's efforts, and thus that there may be a tacit internal contradiction in his views. Clifford takes careful measure to repeatedly emphasise the dangers and ill-consequences awaiting us in irrational belief, in spite of his putative rejection of those consequences as being of any significance to the ethical character of belief. Does he *really* believe what his philosophical system upholds, or does he simply employ it in a pragmatic attempt to spread rationalist sentiments?

Alvin Plantinga has objected to Clifford's thesis in a number of ways, firstly by questioning whether belief volunteerism is true. Are we in fact able to believe otherwise than we do? Do we possess any meaningful freedom of belief? He has suggested that, aside from extraordinary indirect measures, it does not seem that we possess much freedom in our believing. Were someone to offer him $1 million to stop believing that Mars is smaller than Venus, he says that he could not collect, for his belief that Mars is the smaller of the two is effectively outside of his control.[205] If we are not free to choose what we believe it can be argued that there simply cannot be any true ethical infraction in our believing improperly, for it would be usually understood that to say that we *ought* to believe in some way or another is to say that we *can* believe in some way or another – but

[204] Ibid.
[205] Alvin Plantinga (2002).

the latter is denied when the free-belief thesis is denied. Plantinga also questions evidentialism – the idea that we require evidence for all that we believe. He notes, as others have,[206] that we simply cannot be expected to have reasons for *everything* we believe, for this leads us into an infinite task of justification that we can never succeed at.

While it may be generally immensely difficult or even impossible to immediately believe that one planet is larger than another by our decision to alone, it is also true that beliefs on matters such as the relative sizes of planets do not constitute or represent all possible types of belief. There are many subject matters in which there seems to be much greater room for consideration and doubt, and therein at least a *prima facie* case for a freedom in belief to be exercised through the modification of considerations and values in deliberation. We do not say, after all, that because a drug addict's choices are severely modified by the neurochemical effects that the drugs have on him, that he therefore has no free will, at all, on *any matter whatsoever*. Although we might certainly deny that we possess something like free will, doing so on the basis of the influences that drugs can have upon us does not provide a compelling basis for that. In the case of planets, our relevant knowledge of their comparative sizes comes to us by testimony: we believe that Mars is smaller than Venus because it is a thoroughly simple and well understood belief of little consequence within a larger body of astronomical knowledge delivered to us by experts in whom we have a high level of confidence. I take it to be fairly obvious that not every belief approaches this same level of intransigence. What is to be said for instance of beliefs we arrive at on our own, from our own experiences and inferences? What of beliefs in which it is possible for self deception to play a role?

Consider the case of the man trying to decide whether his wife is cheating on him or not. He notes that lately she has been regularly arriving home late from work. The husband has no reason to believe that the job she is employed at would demand such irregular working hours, nor is he aware of any particular reasons why her job should only recently demand this added work. He notes that she has been in his judgment a little too friendly with a particular male colleague at functions they have attended together, and he has been acutely aware for some time now that their own relationship has grown distant (I will let the reader imagine numerous other small details in order to establish a plausible case for her infidelity). But in spite of this there is also evidence to the contrary. He notes her as not being 'the sort of person' who would cheat in a relationship, as having both a trusting and trustable nature, and so on (again, I will let the reader imagine other plausible points which would support her innocence). Thus there is room for doubt and for belief in either direction. In this sense, might he be able to *choose* what he believes? Certainly the case for a liberty of belief is much better here than the case of believing contrary to well established basic scientific facts. How he weighs the relevant evidence in his deliberations will affect what he comes to ultimately believe, and he may have some measure of freedom over that.

[206] Popper called a rationalism that does not recognise its own limits regarding evidence - and indeed the necessity of some beliefs to be free of the demands for evidence - "uncritical rationalism", and "comprehensive rationalism".

Suppose again that in fact the evidence against her innocence is substantial indeed, and the evidence in favour of her innocence not great at all. Is the man's belief then forced? Well, perhaps not, for out of his longing desire to avoid being the victim of the pain that would come with the belief in her infidelity, he may engage in a kind of self-deceptive process. He may choose to focus much more upon the evidence that supports her honesty, and spend in turn a much greater amount of time disconfirming evidence in favour of her guilt. Every time a doubt arises in his mind he may put it to rest by dismissing it (with some effort, perhaps, but consistently nonetheless). He might in effect create a sort of illusion for himself, and convince himself that his doubts and fears are the irrational result of his own weaknesses of character. But we would probably allow that he need not have done so. We would think that he *might have* faced the looming truth of the evidence before him, and note in this case that he simply chose not to. He had a choice between the harsh and apparent truth, and a self-imposed and created fantasy of sorts – a choice more easily available to him in that case than for us to alter our beliefs regarding the relative sizes of planetary bodies. As John Rawls has argued, our judgments are subject to various 'burdens' that make determinations difficult and allow for reasonable differences of views between people.[207] For instance, evidence can often be complex and conflicting, it can be difficult to determine what is evidentially relevant, and difficult to know what weight to assign to different points of evidence. From within this difficulty there may well reside room for a freedom of belief. As such there are reasons to doubt Plantinga's argument against doxastic freedom, and for us to think that as an objection to Clifford it is unsuccessful.

Regardless of this, other elements of Clifford's position raise concern. There is an obvious connection between what we believe and how we act, but there are reasons to think that this connection is not straightforward. Let us consider again the theory-ladenness of observation in the case of science, since the argument I will make here takes essentially the same form. As noted by Duhem, scientific theories are not tested in isolation but in the company of many other assumptions, theoretical commitments, initial conditions, and so on. It is only holistically then that theories may be corroborated or falsified by being brought to measure against the way the world is observed to be. This has an analog to action.

Consider a simple case: a belief B by person P that a pen is required now in order to jot down some philosophical notes. Call A the set of physical actions involved in acquiring the pen. Does B lead directly and immediately in and of itself to P's doing A? It does not. P is required to either recall the location of a pen known to be in the vicinity, or one thought to be probably there, or to scan around the room in search of one by chance, all of which involve many beliefs about the way the world is. P is required to tacitly believe that the pen will remain in their hand when grasped between fingers and thumb (a person who did not believe this may not even try to acquire the pen) and have reasons to believe that the pen being searched for, or that is likely/possible to be found, will work or can work (e.g. that it is not likely to be empty of ink or broken in some way that prevents it from a useful employment). We can see then that even a simple task like this relies upon many background beliefs being held.

[207] Simon Clarke (1999), citing John Rawls 'Political Liberalism' (1993).

Let us take a more complex and grizzly scenario in order to highlight the moral element of belief. P2 believes that 'killing innocent people for the pleasure of it is good'; call this belief B2. Does B2 lead to an act or acts of murder alone? It does not. In order for murderous action A2 to transpire as a result of B2, B2 must be conjoined with a number of other beliefs about the way the world is. Our prospecting murderer needs to know many things, such as what kinds of instruments or methods can be used to kill people, under what conditions those instruments will prove deadly, whether they are available to him already or need to be procured, and so on. He needs further basic knowledge, such as how to get to a place where a murder can be committed, which could itself involve a far deeper detailed knowledge of things like transport systems and other practical knowledge. He also likely needs to know much about himself. The 'good' referenced in the belief in question is most likely conditional on various things obtaining, such as not getting caught in the process or thereafter, and not feeling guilt afterwards (in which case that belief will be an approximation of a more detailed belief set involving all those conditions which modify the conditions of the anticipated 'goodness' of the action). Thus, the act made in light of this belief will require a further wealth of information in order to live up to its expectations – namely, that an act will instantiate that good expected. Extrapolating from such cases, what we can see is that it is largely *webs of belief* that collectively produce action. Many of the beliefs involved in this we overlook simply because they are too mundane, widely shared, or natural to explicate, but that is precisely what Duhem criticised in relation to the philosophy of science. We need to be aware of them, he argued, if we are to have an accurate philosophical account of what is going on. In this case, they are philosophically important because their presence or absence influences what actions derive from our beliefs.

What if we consider non-action as a kind of action? Compare the belief that 'it is wrong to kill innocent people' with the belief that 'it is good to kill innocent people'. The first suggests a refraining from action, or non-action, whereas the second suggests a positive action in the world. We *could* consider, although we normally would not, the first as a kind of action itself. But even the act of 'not killing innocent people' requires a great many beliefs about the world – namely, and at a minimum, what constitutes killing/death, what constitutes 'people', and in what ways it can be avoided (i.e. don't drive while intoxicated, don't run with knives, don't play with guns, etc, all of which requires much deeper and richer sets of beliefs in turn). Thus this refraining from action is still subject to the same belief embeddedness that draws extensively upon our background knowledge. The point is that the connection between belief and action is not so straightforward as might first be thought, and that as inputs into our mental networks beliefs need not go ineluctably towards bringing about any actions we might associate with them. The connection between belief and action is far more complex than simple linear associations allow for, depending in many cases upon other beliefs before being translated into actionable results.

This essential point can be further strengthened when we consider that beliefs are best thought of not in terms of their being held at all, but in terms of the strength to which they are held. One may believe that astronauts never landed on the moon (but only as a

hunch). One may believe that astronauts never landed on the moon (and be utterly convinced of this). Both of these beliefs are logically the same (belief in ~x), yet both are clearly different from one another in important respects. If for instance I believe that god exists (but with some doubts), the way in which I might approach an associated religious text is likely to be different from the case in which I believe that god exists (in my heart of hearts, with unfailing confidence), when all other things are equal. The mere presence of a belief, absent of information regarding the strength to which it is held, does not tell us with clarity how it will affect or engage with other beliefs concurrently held to produce any associated action(s). Even if in both cases of passion and doubt I were to approach a religious text in an identical way, the acquisition of further additional beliefs later (such as the belief that there is some evidence that the text in question is a man-made creation) could modify my approach differentially in each case.

Clifford would principally reject these points as relevant, since we know that he held the moral infraction of any belief to be contained purely within its rational standing, rather than in any action it produced. Any question then of the distance between a belief and an action is of no significance to his deontological weighing of the ethics of belief. But if we reject this deontic approach the distance between belief and action immediately becomes relevant again, for the 'guilt' of any belief may then reside in the significance of its role in causing any harmful action or states of affairs, in relation to features of its rational character. If the connection between belief and action is straightforward and immediate, then the wrongness of an action can transfer immediately and unimpeded to the irrationality of the belief that caused it. However, if it is only sets of beliefs that cause an action, and if beliefs are held in variable degrees of cognitive strength, then there can easily be some degeneration in the ethical transfer of wrong action to rationally ill-founded belief. Taken together then, these two points, that a) even quite basic beliefs can require others in order to lead to their 'suggested' actions, and that b) the etiological character of beliefs expressed in action becomes vastly more complicated when we allow for variations among the strengths beliefs are held at, drive a small wedge between the immediacy of belief to action, and so undermine naïve views of the role beliefs play in producing actions (which as was stated is relevant to non-deontological assessments).

What can be said then of a private sphere of belief? If the connection between belief and action is less than immediate and certain, there are grounds on which to defend a space of belief in which the interests of others do not reach. If some beliefs are only distantly related to the actions they are taken as contributing towards, then irrespective of their rational standing it may be quite unreasonable to hold them as subject to collectively imposed expectations or standards. My belief in the existence of an invisible pink teapot buried on the northern polar icecap of mars, for instance, may lead to nothing directly actionable, and nothing (to my knowledge) indirectly influential upon my actions. But even if it did, its role in that may be so minimal in comparison to other beliefs as to most prudently escape our attention. For instance, perhaps this belief goes towards shaping my attitude in some minor way, which itself is shaped by countless other beliefs too plentiful to enumerate, and my attitude causes (in conjunction with many other beliefs) a single distinct action that is judged to be mildly wrong. In such a case this particular belief

features so far back and so weakly in the causal chain of events leading to the action that it may not be worthwhile at all to allocate it any attention at all – moral or otherwise.

I don't deny that many beliefs we have are of an interest to those around us. Public education systems are predicated upon society having a legitimate stake in what people believe, and these interests may extend to set what is taught to individuals.[208] But the mere existence of a public interest in what an individual believes does not, in and of itself, make legitimate sweeping demands in that respect. It would be, I think, a quite undesirable turn of events if an individual's private world was held hostage to the interests of society and subsumed under extrinsic duties to others. My intuitions are that there are limits to the claims of others over us with respect to our thinking, but this freedom need not preclude the existence of rational standards within that internal space. As I will argue, rational urgency and commitment can obtain within a person in an autonomous, rather than heteronomous way.

Lastly in relation to Clifford, we can object to his thesis on pragmatic grounds. It seems overly provocative and unnecessary to demand that on the issue of each and every belief we are duty-bound to rational standards for the good of others. It would make us all morally guilty, for none of us can ever examine enough of our beliefs, and enough of them sufficiently thoroughly, to actually fulfill what is being demanded by moral duty. We simply do not have the mental, energetic, or economic resources to engage in so dedicated an action as this. Clifford's only response to the pragmatic constraints of rational inquiry is dismissive. He chides those who cannot find the time to adequately inspect their beliefs in the manner he sees as necessary should "have no time to believe [at all]".[209] This is clearly an inadequate response. Further to the point, inducing angst in people on every issue of belief, I suggest, is not the best way towards cultivating sentiments for rational thinking in people. We do not normally perceive that each and every instance of belief presents a moral crisis – and making this out to be the case is unlikely to curry favour widely. Nor may it be prudent to frame the issue in deontological terms. Voltaire, in a letter written while exiled in Tulle for writing poetry the authorities found objectionable, wrote,

> I am at a chateau, a place that would be the most agreeable in the world if I had not been exiled to it, and where there is nothing wanting for my perfect happiness except the liberty of leaving. It would be delicious to remain, if I only were allowed to go.[210]

His sentiments capture what I think underlies to some extent a tendency of thinking present in many, including myself. It is one that seeks to resist being duty-bound, and instead favours thought and action which inwardly results. In my own experience, and in my experience of others, I have found people to be generally more willing to act in some way and not another if they are only given a choice and are asked to judge for

[208] Alvin Goldman (1999), pp. 350-351.
[209] William Clifford (1877).
[210] Robert Ingersoll (1894).

themselves; if they are engaged as agents, and thus respected in a certain basic way, rather than dictated to by the demands of putative duties.

If so, it is an advantage of any appeal to rationality that it avoids attempting to strong-arm rationalist conformity. The central reason for our being rational can, alternatively, take the form not of our being externally 'pulled' towards rationality, but 'pushed' endogenously in that direction by ingrained, pre-existing (and perhaps largely inescapable) rational principles and values.

5.3 Promoting Autonomous Rationality

My appeal here proceeds by presupposing a basic 'rational unity of mankind'. It might be charged that it is naïve or overly optimistic to hold out that there is anything like a core rational feature common to (nearly) all, but we can find a perfectly consistent underlying rationale or internal logic to many ways of thinking and behaving that we nevertheless consider to be 'irrational'. Rationality is in some sense a relational term, for an act, argument, or belief is only rational *in relation* to some state or goal. A line of reasoning is said to be rational if it is formed in a way such that the truth of the premise(s) supports the truth of the conclusion(s), and irrational if it fails to satisfy that condition. This is so because arguments characteristically deal in the truth status of propositions. Similarly, we most commonly think that beliefs are rational if they are evidentially well supported, as our default aim with belief is to have true beliefs (and evidence is truth-supporting). But we need not take having true beliefs to be the only possible aim of belief. That is, we *could have* alternative aims that preclude true-belief aims. If my sole aim in believing is to have pleasant beliefs, avoiding being exposed to evidence that conflicts with what I preferentially believe bears out a quite rational character within its own confines, in contradistinction to the normal veridical end we seek. If one's aim in providing others with argument is not to establish truth in rationally legitimate ways, but to swindle people into believing certain things for ones own advantage, then any convincing but heretofore irrational means of arguing carries its own 'rationality' other than the truth-considered. And in the case of actions, if my aim is to get to work on time when I am late, driving slowly is, all other things being equal, irrational; however, if my aim is to lose my job, that action is perfectly rational insofar as it promotes that end. Many of those who do not exhibit rational behaviour or beliefs from the point of view of, say, true or traditionally reasonable beliefs, may still exhibit in their own way an antecedent rational process that bears out a logic. This rational feature may not be the same as the full blooded sense of rationality we by default think of in considerations of the rational merit of something, but it does reveal an underlying rational framework and process. So there is perhaps nearly always a more basic sense of rationality to found.

But we can say more than this. While behaviour or beliefs may be rational from within a certain frame of reference, I suspect that it often cannot be helped but to utilise that competing and more orthodox framework of rationality that takes truth as its aim, and so not really possible to escape that sense of rationality consistently or entirely. To recall, the problem with irrationalism according to Popper is not that it is completely irrational,

but *inconsistently* rational. Yet so long as some meaningful range of a person's beliefs is held on a reason-and-evidence orientated basis, all beliefs are principally under the threat of examination on those very same grounds. Compartmentalisation, the idea that we can internally have quite separate standards of belief in relation to different topics, may work *in effect* to close off irrational beliefs from rational inspection, but it does not seem necessary that this is so, or necessarily consistently so. Thus if a person has rational characteristics in the sense favourably endorsed here, it may be quite possible to appeal to those standards insofar as they *are* applied in order to urge a more universal application. How successful any such appeal can be is an open question; at the end of the day there can be no guarantee of success, for it is the notably resilient character of irrationalism that is the cause of much rationalist frustration. In any case, the deeper interest here is not in stamping out irrationalism entirely, but in extending rational activity and thought in those who are already principally recognising of it.

There might be diverse ways of rearranging social institutions and altering social practices open to us in order to maximise rational choice and to contribute towards less biased and more considered judgments in people on matters of importance. For instance, in cases where we understand a task to be highly important to ourselves or to others, or where we are led to believe that we will be expected to defend what judgments we arrive at, we tend to be more careful and select more consuming or complex methods for coming to considered judgments. Additionally, where subjects expect to be evaluated, to justify their decisions or determinations, to have their judgments made public, or understand their judgments to be of some meaningful consequence to an evaluated person's life, they tend to show less primacy effect in impression formation and employ fewer stereotypes in their evaluations.[211] Inferring from this we can see that there may be ways in which social settings can be manipulated to allow for maximal rational choice in agents. In a similar fashion we may appreciably note that many features of liberal societies *already* beneficially contribute towards its appreciable epistemic condition, and so seek to support those features insofar as they do not exist for illiberal societies. Allen Buchanan has argued that several key features of modern democratic societies are conducive to the reduction of moral and prudential risks that are caused by false beliefs, through making possible and efficient the discovery of error. Those features include freedom of thought, conscience, expression and association, the role of merit in the identification of reliable sources of belief, the level of epistemic egalitarianism that exists within such a society, and the promotion of various resources that serve to examine the social processes of belief formation and deference.[212] The more developed such features are within a society, the better the epistemic situation there is, all other things being equal. It is thus important to keep in mind that while my focus is upon broad attempts to positively influence the epistemic character of individuals, and society as a whole, it should not be thought that other considerations, such as those just outlined, are not worth pursuing towards that end.

What I argue for, I confess, is conventional. It is that we should attempt to instill in people through various education channels and means an understanding of our delicate

[211] Ziva Kunda (1990).
[212] Allen Buchanan (2004).

and vulnerable epistemic situation, socially and cognitively. Its modesty or conservatism can be seen as a distinct advantage however if we grant some element of truth to Popper's warnings regarding revolutionary attempts to solve worldly problems, that "the attempt to make heaven on earth invariably produces hell".[213] What seems better called for by prudence and wisdom in the light of history is not utopian dreams and methods, and their accompanying hoo-ha and claims of deliverance, but piecemeal attempts to better our world by ever watchful increments. Education is I take it, of all things, a relatively benign means of rational and epistemic improvement, yet significantly promising in its own way. I see it as applying *in conjunction with* other possible means of improving our epistemic condition, and to some extent presupposing a basic level of social development in which a project of its kind can be viewed as possible and worthwhile at all.

The benefits of this in socially favourable conditions are clear. Once we as a society are informed in greater detail of the epistemic picture of things, an inward found sense of rational importance is likely to develop quite naturally on its own, for once we see how important the truth of our beliefs is, we see at once the value of adopting measures to aid to ensure their truth. We are in such cases internally pressed towards greater levels of scrutiny as a direct response to awareness of how the truth of our beliefs is at many points undermined. Just as how the sight of a small bird found fallen on the road from its nest and standing little chance of survival may provoke within us a concern for its welfare, and benevolent behaviour towards it on that basis (rather than from any 'duties' that it might be claimed that we owe to it), so too can a drive to better our epistemic condition develop inwardly once having been set alight by understanding in greater clarity and availability just how precarious and open to error that situation is. It all follows out of a basic concern for the truth of our beliefs: that it is desirable to understand and sharpen our critical inquiry in order to eliminate or reduce irrationality, bias, prejudice, and error (all fully basic, fundamental rationalist values). The strength of this desire depends upon the extent to which we recognise ourselves and others as being susceptible to those epistemic dangers, hence the value of bringing to light that reality through salient examples.[214]

Providing people with the skills and knowledge necessary to analyse and better understand social and cognitive epistemic situations will involve providing information regarding the many areas of fallibility facing us, including social and cultural forces and pressures, the historical, theoretical, and linguistic characters of expertise, and the veridical, moral, and prudential risks presented therein, along with the necessary and essential value of the epistemic division of labour and epistemic dependence upon others. Yet the importance of cultural and social institutions in transmitting beliefs, values, skills, and attitudes must not be lost in consideration, and should be given substantial emphasis in order to avoid *radical* scepticism in that regard. What should be sought after is a cultural and social atmosphere that fosters a critical attitude which preserves and emphasises the individual's right and place of questioning sources of belief, along with the broader epistemic and social value of this. The end product is a community that extols

[213] Karl Popper (1962), vol. 2. p. 262

[214] I thus turn again to my defence in the first chapter regarding the merits of understanding how prominent historical thinkers and figures failed to strike truth in their beliefs in every case.

the virtue of critical reasoning and standards, while denying culture as having any special, immutable authority.

To be clear, and to pre-empt a possible criticism: there is no tension in the fact of our beliefs and values about proper ways of epistemically positioning ourselves with respect to beliefs and values, and in particular about cultural and social sources of belief, arriving in us through cultural and social means of transmission, for a critical attitude encouraged and made available in this way immediately properly becomes its own object of examination. In other words, if it is taught that it is good to be critically demanding about what one is taught, then one is taught (implicitly, but also advisably explicitly) to be critically demanding of the *very idea* of being critically demanding, for this is something itself taught to us. Critical thinking, then, must pass its own test.

5.4 The Two-Fold Root of Epistemic Egalitarianism

Optimized epistemic egalitarianism is both beneficial to and bound up in the quality of the social epistemic condition, for the conditions that it specifies are *part of what it would be* for a society to be epistemically well-formed. Epistemic egalitarianism, the state in which people feel sufficiently confident enough in themselves to ask questions of authorities in the case that they see reason enough to do so, is not something that a society either strictly lacks or possesses, but rather something that it bears out to greater or lesser extents. It does not specify a culture of questioning, but rather a proper or desirable balance between trust and distrust of social suppliers of belief, in relation to the reasons available for each. It is essential then to note that the quality of epistemic egalitarianism within a society can be promoted independently of *acts* of questioning, for a person's ability to question authority depends in a large part upon the relevant level of background knowledge that enables them to at all. Accordingly, promotion of background knowledge relevant to epistemic issues can be seen as positive *even if* the epistemic situation of a society is in fact judged to be in such good order that individuals there find little reason to do so on most occasions. As such one important measure of the epistemic health of a society is its *capacity* to effectively question putative authorities, independently of instances of questioning (which will depend upon judgments related to any cases thought to be requiring that).

There are two central reasons which can serve as a fundamental basis and motivation for our seeking to promote epistemic egalitarianism. They do not constitute the only reasons for which it might be sought after, but they are treated here as having principle significance in that respect:

1) Recognition in light of psychological and social facts about us of the possibility of bias and error in testimony and authority. A high level of intelligence or a significant social standing is no assured defence against error and its associated consequences. Experts are subject to many of the same sorts of biases and subjective dispositions that can impair epistemic and moral faculties generally, yet face some additional ones that are unique or particular to the status of

expertise itself. Institutional arrangements and self-critical procedures may fail to recognise and overcome them. One is therefore principally entitled and normatively commended to approach sources of authority with a critical posture *a priori*.

2) Recognition in view of history and contemporaneously of the moral and prudential danger that false belief presents, particularly when those beliefs are dispersed widely through epistemic channels and integrated into the social whole. Many of those recognised as having been in possession of profound intelligence and extensive learning, and who have been afforded a deferential status by members of society as a result, have simultaneously endorsed what we may rightly view as seriously immoral or dangerous ideas, towards whom acts of deference on those matters in question would have been imperilling. One is therefore principally entitled and normatively commended to approach claimant sources of authority with a critical posture *a posteriori*.

Considerations must also extend towards the political, legal, and social freedom of individuals to question authority, independently of their knowledge-dependent capacity to do so, and their willingness to. If those who are epistemically enabled to effectively question social or political decisions nonetheless fail to do so out of justified fears for their own well being, such as through the threat of severe governmental retribution, then it is clear that there are more primitive limiting factors at fault for the degraded state of epistemic egalitarianism in that situation. Accordingly, advancing levels of relevant knowledge is by no means the only way towards advancing epistemic egalitarianism, as other social features must be in place in order for it to be fruitfully exercised.

5.5 On a Place for Critical Thinking within Education

The traditional aim of education is most characteristically a veritistic one. That is, public education systems attempt to provide individuals with knowledge of a certain relevant sort.[215] It might be thought that this end implies that the best means towards achieving this is to teach in the manner of direct tellings of recognised facts (rote learning, the 'stand and deliver' method). This would be a mistaken inference to draw, however. Let us say that education succeeds to the extent in which individuals acquire significant new beliefs which are true, or to the extent that it increases their degree of belief in such beliefs. A conception of education with truth as its central aim is conceptually neutral in regards to how this might be best achieved. Along side basic aims of transmitting pre-existing significant knowledge, and providing incentives to learn, education systems may well seek to facilitate an uptake of intellectual skills, values, and background beliefs that enhance an individual's independent learning and reasoning skills. Hence, not only is

[215] Alvin Goldman (1999), p. 349. Goldman distinguishes between various interests in what is taught – i.e. what is of interest to the student now, what *would* be of interest to the student if they knew more, and what it is in the interests of the community that individuals belonging to it know. Presumably there are a virtually infinite number of completely uninteresting, irrelevant facts that education systems have no interest in teaching, or people in knowing (e.g. the number of blades of grass in a given backyard).

education of the sort advocated desirable in light of veritistic goals (irrespective of whether one considers critical thinking to be an end in itself, or a means to an end), [216] but it is not at all at conceptual odds with the traditional, enlightenment conception of the role and function of education within society. What is advocated then fits naturally with pre-existing education goals.

As it stands, in a number of countries education curriculum have begun placing more emphasis upon the importance of critical thinking skills in education. Studies on the effects that philosophy can have on children, such as in the form of critical reasoning skills and cooperative discussion, are encouraging.[217] For instance, philosophical classes have been seen to promote reasons-giving for judgments arrived at by children and promote self-esteem and good class behaviour as compared with control groups.

[216] Ibid, pp. 362-363
[217] See S. Trickey and K. J. Topping (2004).

Conclusion

Given existing commitments to rationality and valuing of the truth, it follows in a basic way that individually and collectively we ought to build upon our background knowledge as it relates to reasoning and thinking in a critical way, especially concerning cognitive and social sources of error. It has been argued that measures be taken to develop a critical attitude that makes widely possible the subjection of the diverse values, beliefs, and attitudes acquired from society to critical inspection, along with the knowledge and skills necessary to better understand our thinking and the way in which it is socially and cognitively operated upon. The method favoured for this is the inclusion in educational programs of the appropriate material necessary to facilitate these goals.

I do not suggest that this is the only method worth pursuing towards that end, only that it is one way of broadly improving our epistemic condition. Education offers a real opportunity for epistemic improvement, for as it is in many cases when we know better, we can very often do better. Through understanding we are afforded a greater power of awareness and greater ability to minimise or overcome the epistemic issues that face us. The interplay between cognitive and social phenomenon creates a massive diversity of epistemic possibilities, towards which a society does well to educate itself in regards to. Within these two broad and overlapping categories, much is contained. To illustrate, under them falls the following sorts of questions:

- In what ways can social expectations and conceptions of social roles influence beliefs and attitudes?
- In what ways can changing social conditions influence beliefs and attitudes?
- In what ways can social reward and punishment systems encourage undesirable epistemic states of affairs? (e.g. doctors and medical paternalism).
- In what ways does the difficulty of identifying experts exacerbate existing epistemic problems?
- In what ways can language influence our tendency to defer to putative experts?
- In what ways can epistemic values instilled in us shape our capacity for critical reflection? (e.g. if a religious community upholds faith as a supreme virtue and hallmark of the pious, and expounds that faith will be rewarded greatly at the end of time, how might this effect people's ability and motivation to think critically and effectively about religious content? If reason is explicitly shunned by prominent religious texts or figures (e.g. Martin Luther), what influences may this have?).
- How can authoritative backing influence the way in which we approach certain claims? (e.g. how does a proposition put to us as a 'scientific fact' affect our stance towards it?)
- How do personal and collective interests (psychological, economic, ideological, political) shape the conclusions that are arrived at by individuals and groups?

Just to *begin* asking these sorts of deeper questions is valuable, even if few concrete or firm answers are arrived at in the short term, because it indicates that the epistemic terrain is understood as presenting problems, evidences an attitude that is willing to ask questions, and shows an exercise of critical reasoning which can be in turn be positively encouraged.

We should seek to teach not simply the facts of our epistemic situation, but our best methods for diminishing the effects of bias and error. For example, in teaching people of the role confirmation bias plays in our reasoning, it would be useful to suggest that one way towards reducing its effects is to attempt to consider alternative hypotheses to those that we favour/are before us in examination. By making cognitive or behavioural strategies available, we can hope to more effectively enable people to think and behave in more clearly rational ways. There is as such much room here for contributions from many areas of expertise, ranging from psychology and social psychology, sociology, to various areas of philosophy, such as epistemology and social epistemology, in order to develop a set of measures to most efficaciously improve upon our epistemic situation. Is education enough? It is highly unlikely to be completely sufficient itself, but what *would be* enough is a question beyond what this thesis attempts, which is not the final solution to problems of irrationality but meaningful steps towards their vitiation.

In that regard, education is advocated as a primary and important step. If I understand that harmful false beliefs have been disseminated historically through existing social roles and channels, I can move from that fact to the possibility of it in any given case, including my own. If I understand the importance that beliefs play in forming moral judgments, as highlighted by the role that false beliefs have historically played in supporting highly unethical positions, I see the importance of true beliefs in my own way of thinking on matters, and in that of other people. When we are brought into awareness of the socially situated, embedded nature of our thinking, of widespread tendencies towards bias, and so on, we can work more effectively towards overcoming them simply as a matter of our rational character and commitments. When we understand and gain a better appreciation for just how reliant we are upon others for what we believe, to what extent we are consciously and subconsciously influenced by those around us with respect to our judgments, beliefs, and attitudes, and how our own preferences, background beliefs, and tendencies in thought shape our reasoning, we gain useful insights which both motivate and enable us to better our epistemic condition, individually and collectively.

Bibliography

Andre, Claire, and Manuel Velasquez. (1990). "The Just World Theory". *Issues in Ethics*, 3. <http://www.scu.edu/ethics/publications/iie/v3n2/justworld.html>.

Audi, Robert. (2004). "The A Priori Authority of Testimony". *Philosophical Issues*, 14. pp. 18-34.

Blackburn, Simon. (2005). *Truth: A Guide*. London: Penguin Books.

Blass, Thomas. *Memorable Quotes*. Dr Thomas Blass Presents: Stanley Milgram.com. <http://stanleymilgram.com>.

Boyd, Robert, and Peter Richerson. (2005). *Not By Genes Alone: How Culture Transformed Human Evolution*. Chicago: The University of Chicago Press.

Buchanan, Allen. (2002). "Social Moral Epistemology". *Social Philosophy and Policy*, 19. pp. 126-152.

----------. (2004). "Political Liberalism and Social Epistemology". *Philosophy & Public Affairs*, 32. pp. 95-130.

----------. (2007). "Institutions, Beliefs, and Ethics: Eugenics as a Case Study". *The Journal of Political Philosophy*, 15. pp. 22-45.

Carroll, Robert T. (2005). "Self-Deception". *The Skeptic's Dictionary*. <http://skepdic.com/selfdeception.html>.

Clarke, Simon. (1999). "Contractarianism, Liberal Neutrality, and Epistemology". *Political Studies*, 47. pp. 627-642.

Clifford, William K. (1877). "The Ethics of Belief". *Contemporary Review*. Reproduced at Secular Web Library. The Ethics of Belief. <http://www.infidels.org/library/historical/w_k_clifford/ethics_of_belief.html>.

Chomsky, Noam. (1995). "Rationality/Science". *Z Papers Special Issue*. <http://www.chomsky.info/articles/1995----02.htm>.

Curd, Martin, and J. A. Cover. (1998). *Philosophy of Science: the Central Issues*. New York: W. W. Norton & Company.

Dawkins, Richard. (1998). “Postmodernism Disrobed”. *Nature*, 394. pp. 141-143. Reproduced at <http://richarddawkins.net/article,824,Postmodernism-Disrobed,Richard-Dawkins-Nature>.

----------. (2006). *The God Delusion*. Sydney: Bantam Press.

Duhem, Pierre. (1954). “Physical Theory and Experiment”. In Martin Curd and J. A. Cover. (1998).

Dutton, Denis. (1995). “Language Crimes: A Lesson in How Not to Write, Courtesy of the Professoriate”. *The Wall Street Journal*. Reproduced at Denis Dutton. Denis Dutton on Bad Writing. <http://denisdutton.com/language_crimes.htm>.

Eagly, A. H., R. D. Ashmore, M.G. Makhijani, and L.C. Longo. (1991). “What is Beautiful is Good, But…: a Meta-Analytic Review of the Research on the Physical Attractiveness Stereotype”. *Psychological Bulletin*, 101. pp. 109-128.

Edwards, K, and E. E. Smith. (1996). “A Disconfirmation Bias in the Evaluation of Arguments”. *Journal of Personality and Social Psychology*, 71. pp. 5-24.

Faulkner, Paul. (1998). “David Hume’s Reductionist Epistemology of Testimony”. *Pacific Philosophical Quarterly*, 79. pp. 302-313.

----------. (2000). “The Social Character of Testimonial Knowledge”. *The Journal of Philosophy*, 97. pp. 581-601.

----------. (2002). “On the Rationality of Our Response to Testimony”. *Synthese* 131. pp. 353-370.

Feyerabend, Paul. (1978). *Against Method*. Verso 4th Ed. Wiltshire: Verso.

Gilovich, Thomas. (1991). *How We Know What Isn’t So: The Fallibility of Human Reason in Everyday life*. New York: The Free Press

Goldman, Alvin. (1999). *Knowledge in a Social World*. New York: Oxford University Press.

----------. (2001). “Experts: Which Ones Should You Trust?”. *Philosophy and Phenomenolgical Research* 63. pp. 85-110.

Gottschall, Johnathon. (2005). “The Heroine With a Thousand Faces: Universal Trends in the Characterization of Female Folk Take Protagonists”. *Evolutionary Psychology* 3. pp. 85-103.

Graham, Peter J. (2000). “The Reliability of Testimony”. *Philosophy and Phenomenological Research* 61. pp. 695-709.

----------. (2006). “Testimonial Justification: Inferential or Non-Inferential?”. *The Philosophical Quarterly*, 56. pp. 84-95.

Guerrini, Anita. (1989). “The Ethics of Animal Experimentation in Seventeenth-Century England”. *Journal of the History of Ideas* 50. pp. 391-407.

Hardwig, John. (1985). “Epistemic Dependence”. *The Journal of Philosophy*, 82. pp. 335-349.

Harris, Sam. (2004). *The End of Faith: Religion, Terror, and the Future of Reason*. Reading: The Free Press.

Harrison, Peter. (1992). Descartes on Animals. *The Philosophical Quarterly*, 42. pp. 219-227.

Hume, David. (1742). **“**Of National Characters”. In *Essays, Moral, Political, and Literary*. 1987. <http://www.econlib.org/Library/LFBooks/Hume/hmMPL.html>

Ingersoll, Robert. (1894). *On Voltaire*. Available at Secular Web Library. On Voltaire. <http://www.infidels.org/library/historical/robert_ingersoll/on_voltaire.html>.

Jones, Karen. (1999). “Second-Hand Moral Knowledge”. *The Journal of Philosophy*, 96. pp. 55-78.

Kanazawa, Satoshi (2003). “Why Productivity Fades with Age: The Crime–Genius Connection”. *Journal of Research in Personality*, 37. pp. 257–272.

Kant, Immanuel. (1960). *Observations on the Feeling of the Beautiful and Sublime*. Trans. John T Goldthwat. NB: University of California Press.

Kitcher, Philip. (1990). “The Division of Cognitive Labour”. *The Journal of Philosophy*, 87. pp. 5-22.

----------. (2001). *Science, Truth, and Democracy*. New York: Oxford University Press.

----------. (2002). “Veridical Value and Social Epistemology”. *Philosophy and Phenomenological Research*, 64. pp. 191-198.

Kuhn, Thomas. (1970a). “Logic of Discovery or Psychology of Research?”. In Martin Curd and J. A. Cover. (1998).

----------. (1970b). “The Nature and Necessity of Scientific Revolutions”. In Martin Curd and J. A. Cover. (1998).

----------. (1977). “Objectivity, Value Judgment, Theory Choice”. In Martin Curd and J. A. Cover. (1998).

Kunda, Ziva. (1990). “The Case for Motivated Reasoning”. *Psychological Bulletin*, 108. pp. 480-498.

Lakatos, Imre. (1977). “Science and Pseudoscience”. In Martin Curd and J. A. Cover. (1998).

----------. *Science and Pseudoscience {transcript)*. Imre Lakatos Memorial Fund, LSE. <http://www.lse.ac.uk/collections/lakatos/scienceAndPseudoscienceTranscript.htm>.

Lester, Greg. (2006). “First Impressions May Demonstrate Why the Pretty Prosper”. Penn University Office of University Communications. <http://www.upenn.edu/pennnews/article.php?id=899>.

Loftus, E. F, and J. C. Palmer. (1974). “Reconstruction of Automobile Destruction”. *Journal of Verbal Learning and Verbal Behaviour*, 13. pp 585-589.

Mencken, H.L. “Memorial Service”. In G. Lee Bowie, Meredith Michaels, and Robert C Solomon. (2000). *Twenty Questions: an Introduction to Philosophy*. 4th Edition. NB: Harcourt College Publishers

Milgram, Stanley. (1963). “Behavioural Study of Obedience”. *Journal of Abnormal and Social Psychology*, 67. pp. 371-378.

Morton, Eric. (2000). “Race and Racism in the Works of David Hume”. *Journal on African Philosophy*, 1. <http://www.africanphilosophy.com/vol1.1/morton.html>.

Naulty, R. A., and P. J. Sheehan. (1975). “Hume, Price, and Testimony”. *Philosophy and Phenomenological Research*, 35. pp. 376-384.

Nickel, Philip. (2001). “Moral Testimony and Its Authority”. *Ethical Theory and Moral Practice*, 4. pp. 253-266.

Nussbaum, Martha. (2000). *The Professor of Parody*. The New Republic Online. Retrieved <http://www.akad.se/Nussbaum.pdf>

Okruhlik, Kathleen. (1994). “Gender and the Biological Sciences”. In Martin Curd and J. A. Cover. (1998).

Pettit, Philip. (2006). “When to Defer to Majority Testimony – And When Not”. *Analysis*, 66. pp. 179-187.

Pinker, Steven. (1994). *The Language Instinct*. New York: William Marrow and Company.

Plantinga, Alvin. (2002). “Theism, Atheism, and Rationality”. *Truth Journal*. Leadership University. <http://www.leaderu.com/truth/3truth02.html>.

Popper, Karl. (1962). *The Open Society and Its Enemies*. Vol. 1 and 2. London: Routledge.

----------. (1963a). *Conjectures and Refutations*. London: Routledge.

----------. (1963b). “Science: Conjectures and Refutations”. In Martin Curd and J. A. Cover. (1998).

Ritcher, Burton. (2006). “Theory in Particle Physics: Theological Speculation versus Practical Knowledge”. *Physics Today.Org*. October 2006, pp. 8. <http://www.physicstoday.org/vol-59/iss-10/p8.html>.

Rosenhan, David L. (1973). “On Being Sane in Insane Places”. *Science*, 179. pp. 250-258.

Rubin, Paul. (2002). *Darwinian Politics: The Evolutionary Origin of Freedom*. N.p.: Rutgers University Press.

Russell, Bertrand. (1915). “The Ethics of War”. *International Journal of Ethics*, 25. pp. 127-142.

Schick, Jr, Theodore, and Lewis Vaughn. (2005). *How to Think about Weird Things: Critical Thinking for a New Age*. 4th Ed. New York: McGaw-Hill.

Schopenhauer, Arthur. (1969). *The World as Will and Representation*. Vol. 1. Trans. E. F. J. Payne. New York: Dover Publications.

----------. *The Essays of Arthur Schopenhauer*. (2004a). eBooks, The University of Adelaide Library. Trans. T. Bailey Saunders. <http://etext.library.adelaide.edu.au/s/schopenhauer/arthur/>.

----------. *Essays of Schopenhauer*. (2004b). eBooks. The University of Adelaide Library. Trans. Mrs Rudolf Dircks. <http://etext.library.adelaide.edu.au/s/schopenhauer/arthur/>.

Shermer, Michael. (2006). “The Political Brain”. *Scientific American*, June 26. <http://www.sciam.com/article.cfm?id=the-political-brain&print=true>.

Soble, Alan. (2003). “Kant and Sexual Perversion”. *Monist*, 86. pp. 55-89.

Sutherland, Stuart. (1992). *Irrationality: The Enemy Within*. London: Constable.

Tajfel, H. (1970). “Experiments in Intergroup Discrimination”. *Scientific American*, 223. pp. 96-102.

Trickey, S., and K. J. Topping. (2004). ‘Philosophy for Children’: a Systematic Review. *Research Papers in Education*, 19. pp. 365-380.

Voltaire. (1733). *Philosophical Letters*. In The Works of Voltaire: A Contemporary Version. Vol. 3. Trans William F. Fleming. Available at the Online Library of Liberty. <http://oll.libertyfund.org/index.php?option=com_staticxt&staticfile=show.php&person=3804>.

----------. (1764). *A Philosophical Dictionary*. In The Works of Voltaire: A Contemporary Version. Vol. 3. Trans William F. Fleming. Available at the Online Library of Liberty. <http://oll.libertyfund.org/index.php?option=com_staticxt&staticfile=show.php%3Ftitle=352&Itemid=27>.